Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 33

Forschungszentrum Jülich GmbH
Jülich Centre for Neutron Science,
Peter Grünberg Institute,
Institute of Complex Systems,
Institute for Advanced Simulation

Lecture Notes of the
43rd IFF Spring School 2012

Manuel Angst, Thomas Brückel, Dieter Richter, Reiner Zorn (Eds.)

Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources

This Spring School was organized
by the Jülich Centre for Neutron Science,
the Peter Grünberg Institute,
the Institute of Complex Systems and
the Institute for Advanced Simulation
of the Forschungszentrum Jülich
on 5 – 16 March 2012

In collaboration with
universities, research institutes and the industry.

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 33

ISSN 1866-1807 ISBN 978-3-89336-759-7

Bibliographic information published by the Deutsche Nationalbibliothek.
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Publisher: Forschungszentrum Jülich GmbH
JCNS, PGI, ICS, IAS
52425 Jülich
Phone +49 (0)2461 61-6048 · Fax +49 (0)2461 61-2410

Cover Design: Grafische Medien, Forschungszentrum Jülich GmbH

Printer: Schloemer und Partner GmbH, Düren

Distributor: Forschungszentrum Jülich GmbH
Zentralbibliothek
52425 Jülich
Phone +49 (0)2461 61-5368 · Fax +49 (0)2461 61-6103
e-mail: zb-publikation@fz-juelich.de
Internet: http://www.fz-juelich.de

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 33

ISSN 1866-1807
ISBN 978-3-89336-759-7

Contents

C07 Polarization Handling: Synchrotron
S. Nandi
C08 The European X-ray Free-Electron Laser Project
M. Altarelli
C09 The Road ahead for the European Spallation Source
C. Carlile

D Techniques

D01 Small Angle Scattering and Large Scale Structures
H. Frielinghaus
D02 Scattering under Grazing Incidence from Surfaces and Interfaces
U. Rücker
D03 Powder and Single Crystal Diffractometry: Chemical and Magnetic Structures
M. Meven
D04 Inelastic Scattering: Lattice, Magnetic and Electronic Excitations
J. Voigt
D05 Atomic Pair Distribution Function: Local Structure Investigations
S. Disch
D06 Critical Scattering at Phase Transitions
K. Nemkovski
D07 Quasielastic Scattering
J. Wuttke
D08 High Resolution Spectroscopy
G. Meier, M. Monkenbusch
D09 Nuclear Resonance Scattering
R. Rüffer
D10 Coherent Imaging with X-ray Free-Electron Lasers
H. N. Chapman
D11 Anomalous and Resonant X-ray Scattering: Spin, Charge and Orbital Order
J. Strempfer

E Topical Applications

E01 Superconductivity
Y. Xiao, Y. Su
E02 Polymers: Structure
W. Pyckhout-Hintzen
E03 Polymerdynamics
D. Richter
E04 Material Systems for Information Technologies
M. Angst
E05 Thin Film Heterostructures
E. Kentzinger, S. Pütter
E06 Magnetic Nanoparticles
A. Feoktystov

Preface

Most of what we know about structure and dynamics of condensed matter systems on an atomic length- and timescale stems from X-ray and neutron scattering. The IFF Spring School 2012 comes timely to the centennial anniversary of the discovery of X-ray scattering from single crystals by Max von Laue, Walter Friedrich and Paul Knipping in 1912. Their breakthrough discovery proved the wave nature of X-rays as well as the microscopic structure of crystals as being composed of periodic arrangements of atoms. In 1914 the Noble prize was awarded to Max von Laue for this discovery. Most of our present-day knowledge on the atomic structure of crystalline and amorphous matter is based on the work following Max von Laue employing laboratory X-ray sources for X-ray crystallography. Since the middle of last century synchrotron radiation with its unique properties was employed for more challenging studies, e.g. in macromolecular crystallography. With the advent of research reactors nearly 40 years later, neutron scattering came into play with its alternate contrast mechanism, its sensitivity to atomic magnetism and collective excitations in solids. Again the Noble prize was awarded to the two pioneers of neutron diffraction and inelastic scattering, Clifford Shull and Bertram Brockhouse, in 1994.

Currently, scattering investigations on condensed matter and life science systems is an extremely rapidly developing field at modern synchrotron radiation sources, free electron lasers, dedicated neutron research reactors and neutron spallation sources. The unique properties of both, synchrotron and neutron radiation, enable groundbreaking research in a very broad range of research topics in physics, chemistry, life science, geoscience, material science and engineering. The range of materials, microscopic structures, phenomena and processes, which can be studied, is nearly unlimited. Experimental methods have been developed and refined, which span an incredible range of length- and timescales from the picometer to meter and from femtoseconds to hours. Doing experiments at these large scale facilities is an especially fascinating and exciting aspect of research for young scientists. Not only will they obtain unique microscopic information of structure and dynamics of matter with innovative techniques and methods. But from the start they will be exposed to cutting-edge technology and work in international collaborations. With the projects of the European X-Ray Free Electron Laser XFEL located at DESY in Hamburg, Germany, and the European Spallation Source ESS in Lund, Sweden, the two world-leading pulsed X-ray and neutron sources will be operated in Europe. Together with the large network of national and international sources in Europe and throughout the world these will provide excellent working conditions for users from universities, research organisations and industry.

Scattering methods at these sources are ideally suited to provide essential and unique contributions to the major challenges facing modern industrial society, such as energy supply, health, environment, transport, and information technology. To this end methods and instrumentation are continuously being further refined. At modern synchrotron radiation sources, beams can be focused down to the nanoscale allowing studies of single nanoelectronic devices. With femtosecond X-ray pulses, ultrafast processes can be monitored at free electron laser sources. A major field of application for X-ray free electron laser sources might become the femtosecond X-ray protein nanocrystallography or coherent imaging of structures of molecules by oversampling techniques. At modern high flux neutron spallation sources, on the other hand, and in combination with deuteration facilities located in proximity of these sources, the dynamics of biological macromolecules can be followed on a wide range of timescales allowing one to visualise the processes of enzymatic reactions. Time-of-flight spectroscopy with polarisation analysis will allow the study of coherent lattice and magnetic excitations in correlated electron systems, which extend up to very high energies and thus make essential contributions to the understanding of electron correlations with the ordering phenomena and excitations of spin, lattice,

charge and orbital degrees of freedom. For both types of radiation, X-ray and neutrons, in-situ studies of time-dependent non-equilibrium phenomena in complex sample environments come more and more into the focus. These few examples taking from the huge abundance of modern applications of scattering methods suffice to demonstrate the enormous current dynamics of the field.

The present course provides a solid introduction into the basics of scattering methods, starting from the interaction processes of electrons, X-rays and neutrons with matter via scattering theory to the application of correlation functions. An overview is given of our current understanding of the structure of crystals and complex fluids, the dynamics of disordered systems and large molecules, as well as collective excitations in crystalline matter. Building on this firm foundation, we then proceed to introduce sources and instrumentation, present-day synchrotron radiation sources and reactor-based neutron sources as well as modern and future pulsed sources, such as free electron lasers and neutron spallation sources. The two European flagship projects, XFEL and ESS, are highlighted. Out of the very lage range of modern scattering techniques, some of the most advanced are presented in detail. In this block of the Spring School, we focus on techniques applicable to topical investigations, such as nanostructures, biological or correlated electron systems. To prepare students for research at these future sources, innovative scattering techniques at pulsed spallation sources or free electron lasers are introduced. Examples of topical applications in various fields of science allow the students to connect to their own research. We point out the application of scattering methods for the grand challenges in information technology, energy materials and life science. Finally, the course is rounded up by a comparison of scattering methods with complementary techniques, such as real space imaging or electron microscopy.

Topics of the lectures include:

- Interaction of X-rays, neutrons and electrons with matter
- Scattering theory and correlation functions
- Collective excitations in crystalline matter
- Dynamics of large molecules
- Sources and instrumentation for synchrotron and neutron radiation
- The European projects for free electron laser XFEL and the spallation source ESS
- Innovative scattering techniques at these future sources,
- Application of scattering to study superconductivity, polymer dynamics, nanostructures, protein crystallography, material systems for information technologies, energy materials etc.
- Complementary techniques

This school could not take place without the help and dedication of many colleagues. We are grateful to all contributors from the Jülich Centre for Neutron Science (JCNS), the Peter Grünberg Institute (PGI), the Institute of Complex Systems (ICS), and the Institute for Advanced Simulation (IAS) from Forschungszentrum Jülich, as well as the colleagues from the RWTH Aachen University as part of the Jülich-Aachen Research Alliance. Explicitly, we acknowledge the time and effort the following colleagues spent to prepare the manuscripts and the lectures:

- ***Angst**, Manuel, Prof., FZ Jülich*
- ***Baumgarten**, Lutz, Dr., FZ Jülich*
- ***Biehl**, Ralf, Dr., FZ Jülich*
- ***Blügel**, Stefan, Prof., FZ Jülich*
- ***Brückel**, Thomas, Prof., FZ Jülich*
- ***Buchenau**, Ulrich, Prof., FZ Jülich*
- ***Dhont**, Jan, Prof., FZ Jülich*
- ***DiVincenzo**, David, Prof., FZ Jülich*
- ***Dunin-Borkowski**, Rafal, Prof., FZ Jülich*
- ***Feoktystov**, Artem, Dr., FZ Jülich*
- ***Frielinghaus**, Henrich, Dr., FZ Jülich*
- ***Gompper**, Gerhard, Prof., FZ Jülich*
- ***Hermann**, Raphael, Prof., FZ Jülich*
- ***Ioffe**, Alexander, Dr., FZ Jülich*
- ***Kemmerling**, Günter, Dr., FZ Jülich*
- ***Kentzinger**, Emmanuel, Dr., FZ Jülich*
- ***Lezaic**, Marjana, Dr., FZ Jülich*
- ***Meier**, Gerhard, Dr., FZ Jülich*
- ***Mokrousov**, Yuriy, Dr., FZ Jülich*
- ***Monkenbusch**, Michael, Dr., FZ Jülich*
- ***Nandi**, Shibabrata, Dr., FZ Jülich*
- ***Nemkovskiy**, Kirill, Dr., FZ Jülich*
- ***Pütter**, Sabine, Dr., FZ Jülich*
- ***Pyckhout-Hintzen**, Wim, Dr., FZ Jülich*
- ***Richter**, Dieter, Prof., FZ Jülich*
- ***Rücker**, Ulrich, Dr., FZ Jülich*
- ***Schmalzl**, Karin, Dr., FZ Jülich*
- ***Schneider**, Claus Michael, Prof., FZ Jülich*
- ***Schrader**, Tobias, Dr., FZ Jülich*
- ***Schweika**, Werner, Dr., FZ Jülich*
- ***Su**, Yixi, Dr., FZ Jülich*
- ***Voigt**, Jörg, Dr., FZ Jülich*
- ***Winkler**, Roland, Prof., FZ Jülich*
- ***Wuttke**, Joachim, Dr., FZ Jülich*
- ***Zorn**, Reiner, Dr., FZ Jülich*

Thanks to Drs. E. Kentzinger, J. Voigt, A. Wischnewski, J. Wuttke and R. Zorn for organizing the rehearsals and coordinating the content of the various lectures.

We are especially glad that several distinguished colleagues from external universities and research laboratories have agreed to contribute to the program of the school:

- ***Altarelli**, Massimo, Prof., X-FEL*
- ***Carlile**, Colin, Prof., ESS*
- ***Chapman**, Henry, Prof., X-FEL*
- ***Disch**, Sabrina, Dr., Université Bruxelles*
- ***Eberhardt**, Wolfgang, Prof., C-FEL*
- ***Klemradt**, Uwe, Prof., RWTH Aachen*
- ***Lehmann**, Eberhard, Dr., PSI*
- ***Meven**, Martin, Dr., TU München*
- ***Roth**, Georg, Prof., RWTH Aachen*
- ***Rüffer**, Rudolf, Dr., ESRF*
- ***Strempfer**, Jörg, Dr., DESY*

Without the participation of all these colleagues, the program would not be as interesting, versatile, and attractive. We would like to express our thanks to all of them for the effort and enthusiasm, which they have put into the preparation and presentation of their lectures and manuscripts. We are very grateful to the board of directors of the Forschungszentrum Jülich for the continuous organizational and financial support, which we have received for the reali-zation of the IFF Spring School and the production of this book of lecture notes. Finally, our special thanks go to M. Beißel and Ms. B. Köppchen for the general management and the help in compiling the Lecture Notes.

Th. Brückel, R. Zorn, M. Angst, and D. Richter February 2012

A 1 100 years of scattering and beyond

Th. Brückel
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 A brief history of x-ray and neutron scattering

"If I have seen further it is by standing on the shoulders of giants.
(Sir Isaac Newton, 1643 - 1727)".

A discovery made exactly 100 years ago revolutionized mankinds' understanding of condensed matter: the observation of interference patterns obtained with x-rays scattered by a single crystal [1]. In 1914 Max von Laue received the Nobel prize in physics for the interpretation of these observations (figure 1).

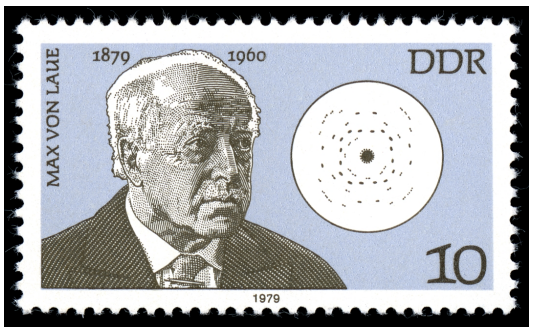


Fig. 1: *Max von Laue and a Laue diffraction pattern on a stamp from the former German Democratic Republic DDR.*

One cannot overestimate the impact of this discovery: it was the first proof that atoms as the elementary building blocks of condensed matter are arranged in a periodic manner within a crystal; at the same time the experiment proved the wave nature of x-rays. The importance of x-ray diffraction for condensed matter research was immediately recognized at the beginning of the 20th century as evidenced by the award of two successive Nobel prizes in physics, one 1914 to Max von Laue "*for his discovery of the diffraction of x-rays by crystals*" and a second one 1915 to William and Lawrence Bragg "*for their services in the analysis of crystal structure by means of x-rays*" [2]. Both, Laue and the Bragg's, could build on earlier experiments by Geiger and Marsden [3, 4] and interpretated by Ernest Rutherford [5] which proved - again by scattering, this time with alpha particles, - that the atom was composed of a nucleus with a diameter in the femtometer (10^{-15} m) range, while the surrounding electron cloud has a typical extension of 1 Å = 0.1 nm = 10^{-10} m. While this seems trivial to us nowadays, this was a breakthrough discovery at the time since alternate models for the atomic structure with a more continuous distribution of positive and negative charges had been discussed and only scattering methods could provide the final proof of the now well accepted structure of the atom consisting of a tiny nucleus and an extended electron cloud. Since these early experiments, a lot of scattering investigations on condensed matter systems have been done. The overwhelming part of our present-day knowledge of the atomic structure of condensed matter is based on x-ray structure investigations, complemented by electron and neutron diffraction. Electrons due to the strong Coulomb interaction with the atoms suffer

multiple scattering events, which make a quantitative evaluation to obtain atomic positions much more difficult. As a probe in condensed matter, electrons made their impact mainly with microscopy techniques (see lecture F 1 by Rafal Dunin-Borkowski). Ernst Ruska was awarded the Nobel prize in physics in 1986 "*for the design of the first electron microscope*".

Entirely new possibilities became apparent with the discovery of the neutrons by James Chadwick [6, 7]. He received the Nobel prize in physics in 1935. However, for neutrons to become a valuable probe in condensed matter research, they had to be available in large quantities as free particles. This was only possible with the advent of nuclear reactors, where nuclear fission is sustained as a chain reaction. The first man-made nuclear reactor, Chicago Pile-1, was built beneath the west stands of Stagg Field, a former squash rackets court of the campus of the University of Chicago. The reactor went critical on December 2, 1942. The experiment was lead by Enrico Fermi, an Italian physicist, who was awarded the Nobel prize in physics in 1938 for this work on transuranium elements. The reactor was a rather crude construction based on a cubical lattice of graphite and uranium oxide blocks. It had no provision for cooling, but two rudimentary manual emergency shutdown systems: one man to cut with an axe a rope, on which a neutron absorbing cadmium rod was suspended, which would drop into the reactor and stop the chain reaction; and a team of three guys standing above the pile ready to flood it with a cadmium salt solution. Considering what was known about nuclear fission at the time it is no wonder that the standing joke among the scientists working there was: *if people could see what we are doing with a million and a half of their dollars, they would think we are crazy. If they knew, why we are doing it, they would know we are [8]*. Out of this very crude first experiment, which was only possible within the World War II Manhattan Project, the modern sophisticated research reactors with their extremely high safety standards developed. In contrast to their big brothers, the nuclear power plants, these reactors are mainly used for isotope production and neutron scattering experiments. The two pioneers of neutron diffraction and inelastic neutron scattering, Clifford G. Shull and Bertram M. Brockhouse, respectively, received the Nobel prize in physics in 1994, many years after the first neutron diffraction experiments, which were performed at Oak Ridge National Lab in 1946. The work of Clifford Shull clearly demonstrated the different contrast mechanisms of neutron scattering compared to x-ray scattering, which in particular allows one to make light elements like hydrogen visible and to distinguish different isotopes like hydrogen and deuterium [9]. But Shull also demonstrated that neutrons, due to their nuclear magnetic moment, could not only be used to determine the arrangement of atoms in solids but they could also be used to determine the magnetic structure e.g. of antiferromagnetic materials [10]. While Shull studied "*where the atoms are located*" and eventually how the magnetic moments are arranged in the solid, Brockhouse observed for the first time "*how the atoms are moving*" in the solid. He developed the so-called triple-axis spectroscopy, which enables the determination of the dispersion relations of lattice vibrations and spin waves [11].

Since the early work in x-ray and neutron scattering sketched above, many years have passed, new radiation sources such as synchrotron radiation sources or neutron spallation sources have been developed, experimental methods and techniques have been refined and the corresponding theoretical concepts developed. For the further development of modern condensed matter research, the availability of these probes to study the structure and

dynamics on a microscopic level is absolutely essential. It comes as no surprise therefore that scattering methods have been employed in ground-breaking work which lead to recent Nobel prizes. The most recent example is the Nobel prize in chemistry 2011, which was awarded to Dan Shechtman "*for the discovery of quasi-crystals*". By means of electron diffraction, Shechtman discovered icosahedral symmetry in aluminum manganese alloys. The observed tenfold symmetry is not compatible with translational symmetry in three dimensions. While the icosahedral symmetry was discovered with electron diffraction, the question where the atoms are located requires the collection of many weak quasicrystal reflections and the analysis of their intensities, which is only possible with x-ray and neutron diffraction. A higher dimensional reciprocal space approach had to be developed to explain the diffraction pattern of such quasicrystals. Another outstanding piece of work in x-ray diffraction is the Nobel prize in chemistry 2009, which was awarded jointly to Venkatraman Ramakrishnan, Thomas A. Steitz and Ada E. Yonath "*for studies of the structure and function of the ribosome*". Ribosomes translate DNA information into life by producing proteins, which in turn control the chemistry in all living organisms. By means of x-ray crystallography the Nobel awardees were able to map the position for each and everyone of the hundreds of thousands of atoms that make up the ribosome. 3D models that show how different antibiotics bind to the ribosome are now used by scientists in order to develop new drugs.

On occasion of the 100 years anniversary of the discovery of x-ray diffraction from single crystals, we have given a brief and absolutely incomplete historical summary of the development of scattering methods. This Spring School is devoted to modern applications of this powerful tool. In this introductory overview, we will now give a short summary of which information we can obtain from scattering experiments, compare the two probes x-rays and neutrons briefly, discuss techniques and applications, giving an outlook into the bright future of the field which the two new European facilities promise: the European X-Ray Free Electron Laser X-FEL (www.xfel.eu) and the European Spallation Source ESS (www.ess-scandinavia.eu) and finally explain how large-scale facilities for this type of research are organized.

2 Introduction to scattering

2.1 Scattering - a critical tool for science

Scattering is the physical process in which radiation or moving particles are being deflected by an object from straight propagation. If the energy or wavelength of the scattered particles or waves, respectively, is the same as before the scattering process, one speaks of elastic scattering or diffraction, otherwise of inelastic scattering. The analysis of the energy of the scattered radiation with respect to the energy of the incident radiation is called spectroscopy.

Nearly all information which we humans as individuals collect on a day-to-day basis about the world in which we live, comes from light scattering and imaging through our eyes. It is only natural that scientists mimic this process of obtaining information in well controlled scattering experiments: they build a source of radiation, direct a beam of towards a sample,

detect the radiation scattered from a sample, i. e. convert the signal into an electronic signal, which they can then treat with computers. In most cases one wants an undisturbed image of the object under investigation and therefore chooses the radiation, so that it does not influence or modify the sample. Scattering is therefore a non-destructive and very gentle method, if the appropriate type of radiation is chosen for the experiment.

What other requirements must the radiation fulfill to be useful for scattering experiments? In condensed matter science we want to go beyond our daily experience and understand the microscopic atomic structure of matter, i. e. we want to find out where the atoms are located inside our samples and also how they move. This cannot be done by light scattering. Why? Well, in general light is scattered from the surface and does not penetrate enough into many materials, such as metals for example. On the other hand, if it penetrates like in the case of glass it is normally just being transmitted except if we have a very bad glass with lots of inhomogeneities. But the main reason is actually that light has a too long wavelength. It is quite intuitive to understand that if we want to measure the distance between the atoms, we need a "ruler" of comparable lengths. The distance between atoms is in the order of 0.1 nm = 10^{-10} m = 0.0000000001 m. Since the distance between atoms is such an important length scale in condensed matter science, it has been given its own unit: 0.1 nm = 1 Ångstrøm = 1 Å. If we compare the wavelength of light with this characteristic length scale, it is 4000 to 7000 times longer and therefore not appropriate to measure distances between atoms. In the electromagnetic spectrum, x-rays have a well adapted wavelength of about 1 Å for studies on such a microscopic scale. They also have a large penetration power as everybody knows from the medical x-ray images.

It should be pointed out that scattering is a much more general method in science, which is not only used by condensed matter scientists. Examples include:

- in the geosciences, seismological studies of the propagation and deflection of elastic waves through the earth are the primary tool for underground exploration (e.g. to detect petroleum bearing formations) and the mapping of the earth's interior.
- the scattering of radar waves is being used e.g. for air traffic control or the detection of weather formations.
- nuclear- and particle physics uses the scattering of high energy elementary particles (electrons, protons etc.) from accelerators to investigate the structure of the nuclei or nucleons etc.

2.2 Scattering cross section

Lets look at a scattering experiment in condensed matter science in the so-called *Fraunhofer- or far-field-approximation*, where we assume that the incident and scattered waves can be described as plane waves with wavelengths λ and λ' (strictly monochromatic) and propagation direction $\hat{\underline{k}}$ and $\hat{\underline{k}}'$, respectively[1]. Let us define the so-called *scattering vector*

[1] Vector quantities are labeled by underlining. Unit vectors are marked by a circumflex ^.

$$\underline{Q} = \underline{k}' - \underline{k} \tag{1}$$

where $\underline{k}$ and $\underline{k}'$ are the wave vectors of the incident and scattered radiation, respectively:

$$\underline{k} = \hat{\underline{k}} \cdot \frac{2\pi}{\lambda};\ \underline{k}' = \hat{\underline{k}}' \cdot \frac{2\pi}{\lambda'} \tag{2}$$

$\hbar\underline{Q}$ represents the momentum transfer during scattering, since according to de Broglie, the momentum of the particle corresponding to the wave with wave vector $\underline{k}$ is given by $\underline{p}=\hbar\underline{k}$. The magnitude of the scattering vector can be calculated from wavelength λ and scattering angle 2θ (between $\underline{k}$' and $\underline{k}$) as follows

$$Q = \left|\underline{Q}\right| = \sqrt{k^2 + k'^2 - 2kk'\cos 2\theta} \Rightarrow Q = \frac{4\pi}{\lambda}\sin\theta \tag{3}$$

A scattering experiment comprises the measurement of the intensity distribution as a function of the scattering vector. The scattered intensity is proportional to the so-called *cross section*, where the proportionality factors arise from the detailed geometry of the experiment. For a definition of the scattering cross section, we refer to Figure 2.

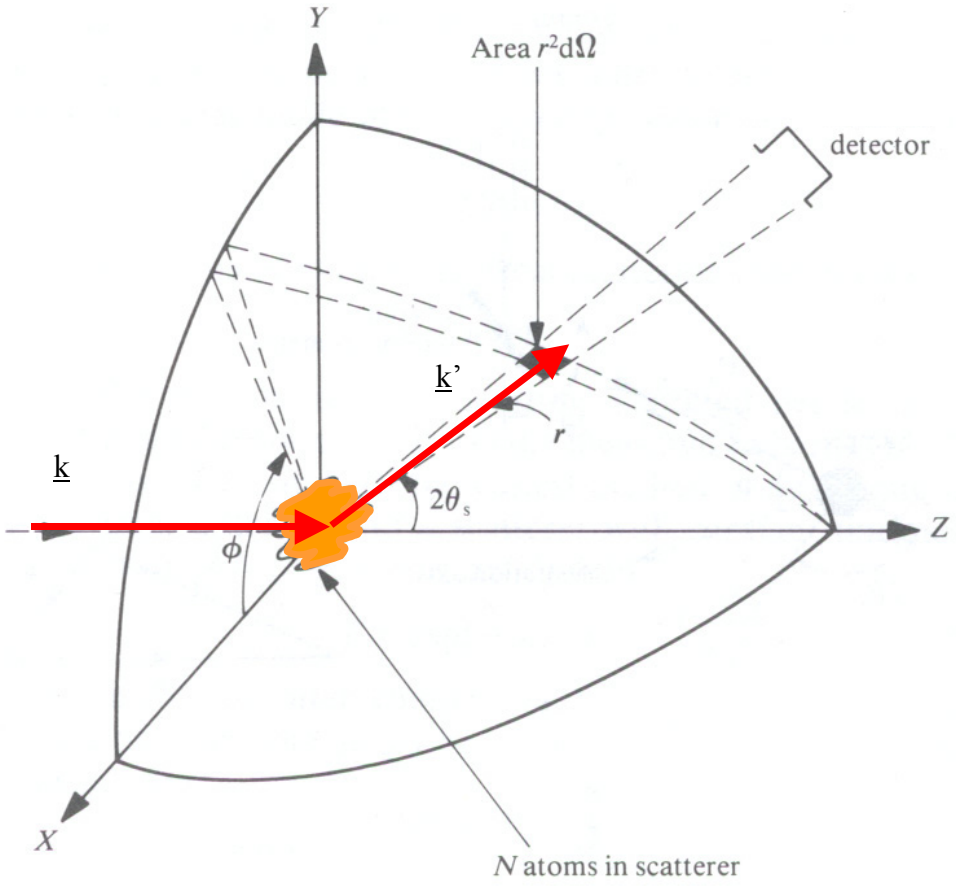


Fig. 2: *Geometry used for the definition of the scattering cross section.*

If n' particles are scattered per second into the solid angle $d\Omega$ seen by the detector under the scattering angle 2θ and into the energy interval between E' and $E' + dE'$, then we can define the so-called *double differential cross section* by:

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{n'}{j d\Omega dE'} \tag{4}$$

Here j refers to the incident beam flux in terms of particles per area and time. If we are not interested in the change of the energy of the radiation during the scattering process, or if our detector is not able to resolve this energy change, then we will describe the angular dependence by the so-called *differential cross section:*

$$\frac{d\sigma}{d\Omega} = \int_0^\infty \frac{d^2\sigma}{d\Omega dE'} dE' \tag{5}$$

Finally the so-called *total scattering cross section* gives us a measure for the total scattering probability independent of changes in energy and scattering angle:

$$\sigma = \int_0^{4\pi} \frac{d\sigma}{d\Omega} d\Omega \tag{6}$$

For a diffraction experiment, our task is to determine the arrangement of the atoms in the sample from the knowledge of the scattering cross section $d\sigma / d\Omega$. The relationship between scattered intensity and the structure of the sample is particularly simple in the so-called *Born approximation*, which is often also referred to as *kinematic scattering approximation* (see lecture A2). In this case, refraction of the beam entering and leaving the sample, multiple scattering events and the extinction of the primary beam due to scattering within the sample are being neglected (these effects will be dealt with in lecture A3). Following Figure 3, the phase difference between a wave scattered at the origin of the coordinate system and at position $\underline{r}$ is given by

$$\Delta\Phi = 2\pi \cdot \frac{\left(\overline{AB} - \overline{CD}\right)}{\lambda} = \underline{k}' \cdot \underline{r} - \underline{k} \cdot \underline{r} = \underline{\underline{Q}} \cdot \underline{r} \tag{7}$$

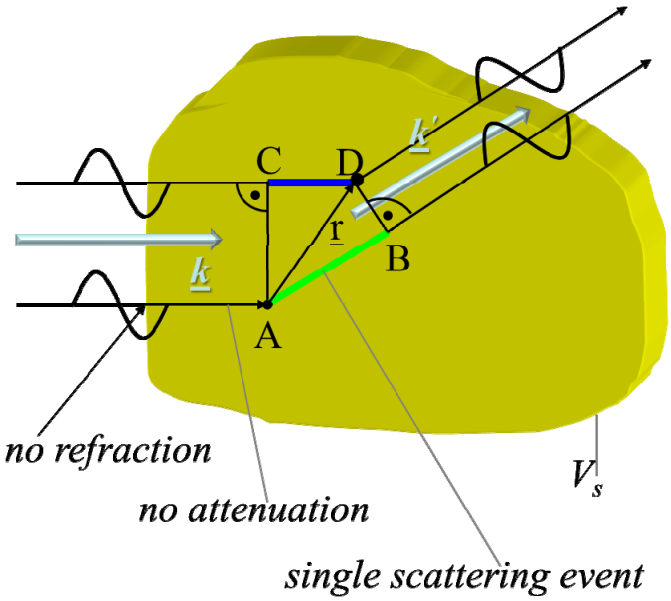


Fig. 3: *A sketch illustrating the phase difference between a beam scattered at the origin of the coordinate system and a beam scattered at the position r. The yellow body represents the sample from which we scatter.*

The probability for a scattering event to occur at position $\underline{r}$ is proportional to the local interaction potential $V(\underline{r})$ between radiation and sample. For a coherent scattering event (interference of scattered waves), the total scattering amplitude is given by a linear superposition of the waves scattered from all points within the sample volume V_s, i.e. by the integral

$$A(\underline{Q}) \sim \int_{V_{\underline{s}}} V(\underline{r}) \cdot e^{i\underline{Q}\cdot\underline{r}} d^3r \tag{8}$$

This equation demonstrates that the scattered amplitude is directly connected to the interaction potential by a simple Fourier transform: *scattering is a probe in reciprocal space, not in direct space and gives direct access to thermodynamic ensemble averages!*

A knowledge of the scattering amplitude for all scattering vectors $\underline{Q}$ allows us to determine via a Fourier transform the interaction potential uniquely. This is the complete information on the sample, which can be obtained by the scattering experiment. Unfortunately nature is not so simple. On one hand, there is the more technical problem that one is unable to determine the scattering cross section for <u>all</u> values of momentum transfer $\hbar\underline{Q}$. The more fundamental problem, however, is given by the fact that normally the amplitude of the scattered wave is not measurable. Instead only the scattered intensity

$$I(Q) \sim \left|A(Q)\right|^2 \tag{9}$$

can be determined. Therefore the phase information is lost and the simple reconstruction of the scattering potential via a Fourier transform is no longer possible. This is the so-called *phase problem* of scattering. There are ways to overcome the phase problem, i.e. by the use of reference waves. Then the potential $V(\underline{r})$ becomes directly accessible. The question, which information we can obtain from a scattering experiment despite the phase problem, will be addressed below and in subsequent lectures.

Which wavelength do we have to choose to obtain the required real space resolution? For information on a length scale L, a phase difference of about $Q \cdot L \approx 2\ \pi$ leads from the primary beam $(Q = 0)$ to the interference maximum. According to (3) $Q \approx 2\pi/\lambda$ for practical scattering angles $(2\theta \sim 60°)$. Combining these two estimates, we end up with the requirement that the wavelength λ has to be in the order of the real space length scale L under investigation. To give an example: with the wavelength in the order of *0.1 nm*, atomic resolution can be achieved in a scattering experiment.

2.3 Coherence

In the above derivation, we assumed plane waves as initial and final states. For a real scattering experiment, this is an unphysical assumption. In the incident beam, a wave packet is produced by collimation (defining the direction of the beam) and monochromatization (defining the wavelength of the incident beam). Neither the direction $\hat{\underline{k}}$, nor the wavelength λ have sharp values but rather have a distribution of finite width about their respective mean values. This wave packet can be described as a superposition of plane waves. As a consequence, the diffraction pattern will be a superposition of patterns for different incident wave vectors $\underline{k}$ and the question arises, which information is lost due to these non-ideal conditions. This *instrumental resolution* is intimately connected with the *coherence* of the beam. Coherence is needed, so that the interference pattern is not significantly destroyed. Coherence requires a phase relationship between the different components of the beam. Two types of coherence can be distinguished.

- *Temporal or longitudinal coherence* due to a wavelength spread.

A measure for the longitudinal coherence is given by the length, on which two components of the beam with largest wavelength difference (λ and $\lambda+\Delta\lambda$) become fully out of phase. According to the following figure, this is the case for $l_{\parallel} = n \cdot \lambda = \left(n - \frac{1}{2}\right)(\lambda + \Delta\lambda)$.

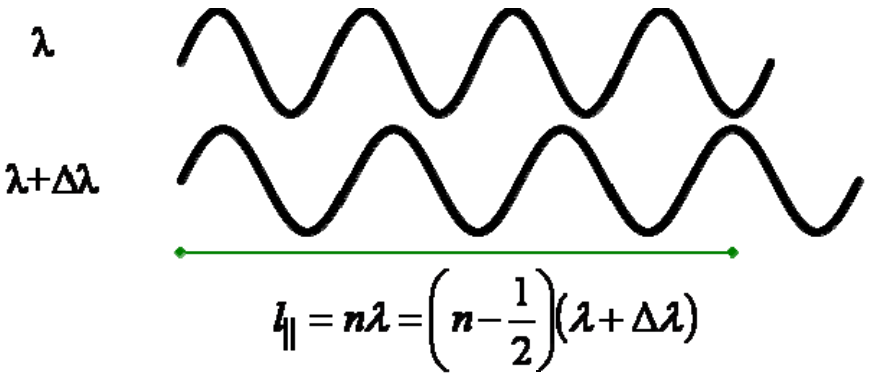


Fig. 4: *A sketch illustrating the longitudinal coherence due to a wavelength spread.*

From this, we obtain the *longitudinal coherence length* $l_{\parallel}$ as

$$l_{\parallel} = \frac{\lambda^2}{2\Delta\lambda} \tag{10}$$

• *Transversal coherence* due to source extension
Due to the extension of the source (transverse beam size), the phase relation is destroyed for large source size or large divergence. According to the following figure, a first minimum occurs for $\frac{\lambda}{2} = d \cdot \sin\theta \approx d \cdot \theta$.

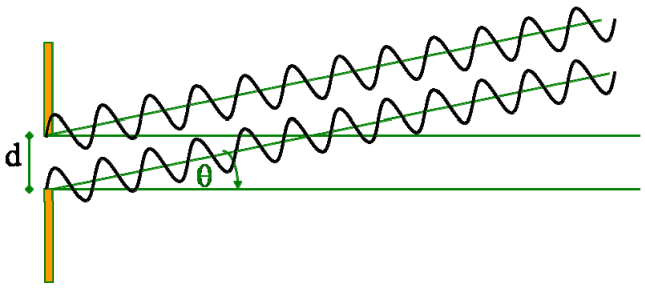


Fig. 5: *A sketch illustrating the transverse coherence due to source extension.*

From this, we obtain the *transversal coherence length* $l_{\perp}$ as

$$l_{\perp} = \frac{\lambda}{2\Delta\theta} \tag{11}$$

Here $\Delta\theta$ is the divergence of the beam. Note that $l_{\perp}$ can be different along different spatial directions: in many instruments, the vertical and horizontal collimations are different.

Together, the longitudinal and the two transversal coherence lengths (in two directions perpendicular to the beam propagation) define a *coherence volume*. This is a measure for a volume within the sample, in which the amplitudes of all scattered waves superimpose to produce an interference pattern. Normally, the coherence volume is significantly smaller than the sample size, typically a few *100 Å* for neutron scattering, up to *μm* for synchrotron

radiation. Scattering between different coherence volumes within the sample is no longer coherent, i. e. instead of the amplitudes, the intensities of the waves contributing to the scattering pattern have to be added. This limits the real space resolution of a scattering experiment to the extension of the coherence volume.

2.4 Pair correlation functions

After having clarified the conditions under which we can expect a coherent scattering process, let us now come back to the question, which information is accessible from the intensity distribution of a scattering experiment. From (9) we see that the phase information is lost during the measurement of the intensity. For this reason the Fourier transform of the scattering potential is not directly accessible in most scattering experiments (note however that phase information can be obtained in certain cases).

Substituting (8) into (9) and applying variable substitution $\underline{R}=\underline{r}'-\underline{r}$, we obtain for the magnitude square of the scattering amplitude, a quantity directly accessible in a diffraction experiment:

$$I(Q) \sim \left| A\left(\underline{Q}\right)\right|^2 \sim \int_{V_S} d^3r' V(\underline{r}') e^{i\underline{Q}\cdot\underline{r}'} \int_{V_S} d^3r V^*(\underline{r}) e^{-i\underline{Q}\cdot\underline{r}} = \iint_{V_S} d^3r' d^3r\, V\left(\underline{r}\right) V^*(\underline{r}') e^{i\underline{Q}\cdot(\underline{r}'-\underline{r})}$$
$$= \iint_{V_S} d^3R d^3r\, V\left(\underline{R}+\underline{r}\right) V^*(\underline{r}) e^{i\underline{Q}\cdot\underline{R}} \qquad (12)$$

This function denotes the so-called *Patterson function* in crystallography or more general the *static pair correlation function:*

$$P\left(\underline{R}\right) = \int_{Vs} d^3r V^*(\underline{r}) V(\underline{r}+\underline{R}) \qquad (13)$$

$P(\underline{R})$ correlates the value of the scattering potential at position $\underline{r}$ with the value at the position $\underline{r}+\underline{R}$, integrated over the entire sample volume V_s. If, averaged over the sample, no correlation exists between the values of the scattering potentials at position $\underline{r}$ and $\underline{r}+\underline{R}$, then the Patterson function $P(\underline{R})$ vanishes. If, however, a periodic arrangement of a pair of atoms exists in the sample with a difference vector $\underline{R}$ between the positions, then the Patterson function will have an extremum for this vector $\underline{R}$. Thus in a periodic arrangement the Patterson function reproduces all the vectors connecting one atom with another atom.

As will be shown in detail in lecture A5, pair correlation functions are being determined quite generally in a scattering experiment. In a coherent inelastic scattering experiment, we measure a cross section proportional to the *scattering law* $S(\underline{Q},\omega)$, which is the Fourier transform with respect to space and time of the spatial and temporal pair correlation function:

$$\frac{d^2\sigma}{d\omega d\Omega} \sim S\left(\underline{Q},\omega\right) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} dt\, e^{-i\omega t} \int_{V_S} d^3r e^{i\underline{Q}\underline{r}} G(\underline{r},t) \qquad (14)$$

While the proportionality factor between the double differential cross section and the scattering law depends on the type of radiation and its specific interaction potential with the

system studied, the spatial and temporal pair correlation function is only a property of the system studied and independent of the probe used:

$$G(\underline{r},t)=\frac{1}{N}\sum_{ij}\int_{Vs}d^3r'\left\langle\delta\left(\underline{r}'-\underline{r}_j(0)\right)\cdot\delta\left(\underline{r}'+\underline{r}-\underline{r}_i(t)\right)\right\rangle=\frac{1}{N}\int_{Vs}d^3r'\left\langle\rho\left(\underline{r}',0\right)\rho\left(\underline{r}'+\underline{r},t\right)\right\rangle \quad (15)$$

Here, the pair correlation function is once expressed as a correlation between the position of N point-like particles (expressed by the delta functions) and once by the correlation between the densities at different positions in the sample for different times. In a magnetic system, radiation is scattered from the atomic magnetic moments, which are vector quantities. Therefore, the scattering law becomes a tensor - the Fourier transform of the *spin pair correlations*:

$$\mathbf{S}^{\alpha\beta}(\underline{Q},\omega)=\frac{1}{2\pi}\sum_{l}\int dt\, e^{i\left[\underline{Q}(\underline{R}_l-\underline{R}_0)-\omega t\right]}\left\langle S_0^{\alpha}(0)S_l^{\beta}(t)\right\rangle \quad (16)$$

α, β denote the Cartesian coordinates *x, y, z;* $\underline{R}_0$ and $\underline{R}_l$ are the spatial coordinates of a reference spin *0* and a spin *l* in the system.

2.5 Scattering from a periodic lattice in three dimensions

We now are ready to understand the famous first diffraction experiment by Laue et al. As an example for the application of (8) and (9), we will now discuss the scattering from a three dimensional lattice of point-like scatterers. As we will see later, this situation corresponds to the scattering of thermal neutrons from a single crystal. More precisely, we will restrict ourselves to the case of a Bravais lattice with one atom at the origin of the unit cell. To each atom we attribute a "scattering length b" (see interaction potential of neutrons below). The single crystal is finite with *N, M* and *P* periods along the basis vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$. The scattering potential, which we have to use in (8) is a sum over δ-functions for all scattering centers:

$$V(\underline{r})=\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{p=0}^{P-1}b\cdot\delta\left(\underline{r}-\left(n\cdot\underline{a}+m\cdot\underline{b}+p\cdot\underline{c}\right)\right) \quad (17)$$

The scattering amplitude is calculated as a Fourier transform:

$$A\left(\underline{Q}\right)\sim b\sum_{n=0}^{N-1}e^{in\underline{Q}\cdot\underline{a}}\sum_{m=0}^{M-1}e^{im\underline{Q}\cdot\underline{b}}\sum_{p=0}^{P-1}e^{ip\underline{Q}\cdot\underline{c}} \quad (18)$$

Summing up the geometrical series, we obtain for the scattered intensity:

$$I\left(\underline{Q}\right)\sim\left|A\left(\underline{Q}\right)\right|^2=\left|b\right|^2\cdot\frac{\sin^2\frac{1}{2}N\underline{Q}\cdot\underline{a}}{\sin^2\frac{1}{2}\underline{Q}\cdot\underline{a}}\cdot\frac{\sin^2\frac{1}{2}M\underline{Q}\cdot\underline{b}}{\sin^2\frac{1}{2}\underline{Q}\cdot\underline{b}}\cdot\frac{\sin^2\frac{1}{2}P\underline{Q}\cdot\underline{c}}{\sin^2\frac{1}{2}\underline{Q}\cdot\underline{c}} \quad (19)$$

The dependence on the scattering vector $\underline{Q}$ is given by the so-called *Laue function* (19), which factorizes according to the three directions in space. One factor along one lattice direction $\underline{a}$ is plotted in Figure 6.

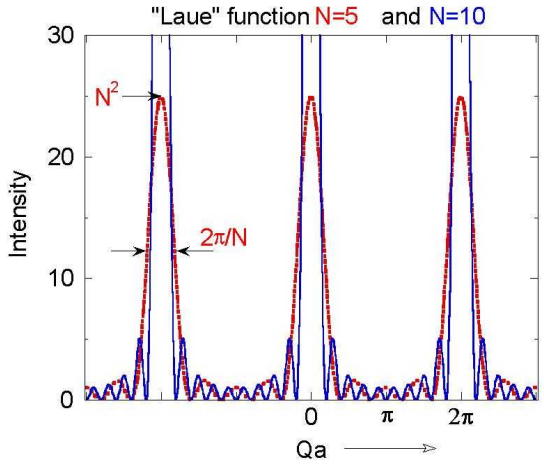


Fig. 6: *Laue function along the lattice direction a for a lattice with five and ten periods, respectively.*

The main maxima occur at the positions $Q = n \cdot 2\pi/a$. The maximum intensity scales with the square of the number of periods N^2, the half width is given approximately by $\Delta Q = 2\pi/(N \cdot a)$. The more periods contribute to coherent scattering, the sharper and higher are the main peaks. Between the main peaks, there are *N-2* side maxima. With increasing number of periods *N*, their intensity becomes rapidly negligible compared to the intensity of the main peaks. The main peaks are of course the well known *Bragg reflections*, which we obtain for scattering from a crystal lattice. From the position of these Bragg peaks in momentum space, the metric of the unit cell can be deduced (lattice constants *a, b, c* and unit cell angles α, β, γ). The width of the Bragg peaks is determined by the size of the coherently scattering volume (parameters *N, M,* and *P*) - and some other factors for real experiments (resolution, mosaic distribution, internal strains, ...).

Via the so-called Ewald construction, it can be shown that the Laue conditions for interference maxima to occur $\underline{Q} \cdot \underline{a} = n \cdot 2\pi$ etc. are equivalent to the *Bragg equation* for scattering from lattice planes (hkl) with interplanar spacings d_{hkl}:

$$2d_{hkl} \sin\theta_{hkl} = \lambda \tag{20}$$

3 X-rays and Neutrons

Since the first scattering experiments, some standard probes for condensed matter research have emerged, which optimally fulfill the requirements for a suitable type of radiation.

First of all, electromagnetic radiation governed by the Maxwell equations can be used. Depending on the resolution requirements, X-rays with wavelength λ about 0.1 nm are being used to achieve atomic resolution, or visible light (λ ~ 350 - 700 nm) is employed to

investigate e. g. colloidal particles in solution. Besides electromagnetic radiation, particle waves can be utilized. It turns out that thermal neutrons with a wavelength $\lambda \sim 0.1$ nm are particularly well adapted to scattering experiments in condensed matter research. Neutrons are governed by the Schrödinger equation of quantum mechanics. An alternative is to use electrons, which for energies of around 100 keV have wavelengths in the order of 0.005 nm. As relativistic particles, they are governed by the Dirac equation. The big drawback of electrons is the strong Coulomb interaction with the electrons in the sample. Therefore neither absorption, nor multiple scattering effects can be neglected. However the abundance of free electrons and the relative ease to produce optical elements makes them very suitable for imaging purposes (electron microscopy). Electrons, but likewise atomic beams, are also very powerful tools for surface science: due to their strong interaction with matter, both types of radiation are very surface sensitive. Low Energy Electron Diffraction LEED and Reflection High Energy Electron Diffraction RHEED are both used for in-situ studies of the crystalline structure during thin film growth, e.g. with Molecular Beam Epitaxy MBE. In what follows we will concentrate on the two probes, which are best suited for bulk studies on an atomic scale: x-rays and neutrons. We will touch upon the radiation sources, briefly discuss the main interaction processes and finally give a comparison of these probes.

3.1 X-ray Sources

Since the early days of Conrad Röntgen X-rays are being produced in the laboratory in sealed vacuum tubes, where electrons from a cathode are accelerated towards the anode. There characteristic- and/or bremsstrahlungradiation is produced. Radiation emitted from such x-ray tubes has been widely used for structural studies in condensed matter science. However, in 1947 a new type of radiation was discovered in a General Electric synchrotron accelerator [12]. It soon turned out that this so-called *synchrotron radiation* has superb properties, see figure 7.

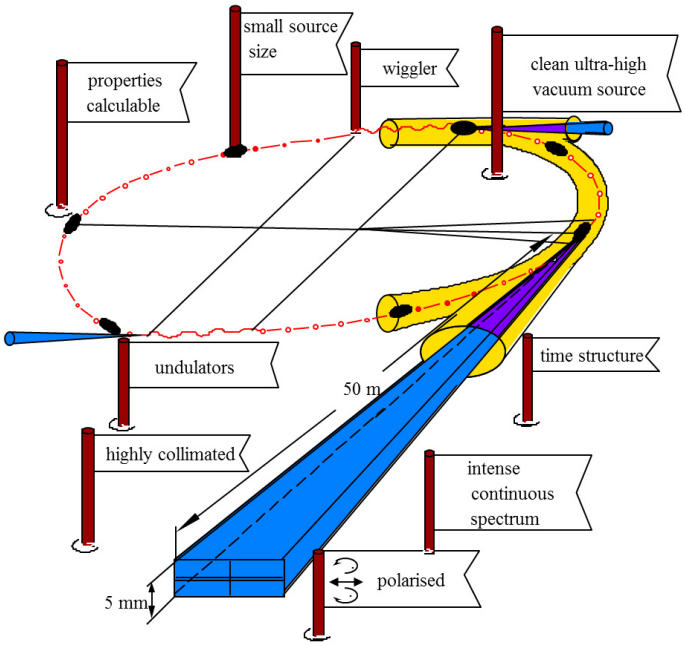


Fig. 7: *Sketch of a synchrotron radiation source indicating the properties of synchrotron radiation.*

Synchrotron radiation is emitted when relativistic charged particles (electrons or positrons) are being accelerated perpendicular to their direction of motion by an appropriate magnetic field. This happens is so-called bending magnets within circular accelerators and this type of radiation has originally been used by solid state physicists in a parasitic mode at particle physics facilities (first generation of synchrotron radiation sources). Second generation synchrotron radiation sources were dedicated to the production of synchrotron radiation, mainly from such bending magnets. However, even more intense radiation can be produced in straight sections of the accelerator by so-called insertion devices - wigglers and undulators - which consist of arrays of magnets with alternating field direction. Modern synchrotron radiation sources of the 3rd generation employ mainly these insertion devices as radiation sources, see lecture C2. This continuous improvement of the source parameters led to an exponential growth of the brilliance, i.e. the spectral photon flux, normalized to the size and divergence of the beam. A further increase of the peak brilliance can be achieved with *free electron lasers*. For the X-ray regime these are based on the SASE principle: *Self Amplified Spontaneous Emission*. In such facilities, an electron beam from a linear accelerator passes through an undulator structure, where synchrotron radiation is produced. The electromagnetic interaction between this radiation and the electron beam travelling in parallel leads to an amplification of the radiation, giving rise to extremely brilliant fully coherent x-ray flashes of about 100 fs duration. Close to DESY in Hamburg such a facility, the *European XFEL* is currently under construction [13]. Details will be presented in lecture C8 by the XFEL managing director, Massimo Altarelli. The facility will open entirely new perspectives for research, see lecture D10 by Henry Chapman.

3.2 Neutron Sources

While neutrons are everywhere - without neutrons we would not exist - they are extremely difficult to produce as free particles, not bound in nuclei. Free neutrons are produced by nuclear physics reactions, which require rather large and high-tech installations. Two main routes to produce free neutrons are being followed today (see figure 8 and lecture C1):

(1) ***Fission*** of the uranium 235 nuclei in a chain reaction; this process happens in *research reactors*.

(2) Bombarding heavy nuclei with high energetic protons; the nuclei are "heated up" when a proton is absorbed and typically 20 - 30 neutrons are being evaporated. This process is called ***spallation*** and requires a *spallation source* with a proton accelerator and a heavy metal target station.

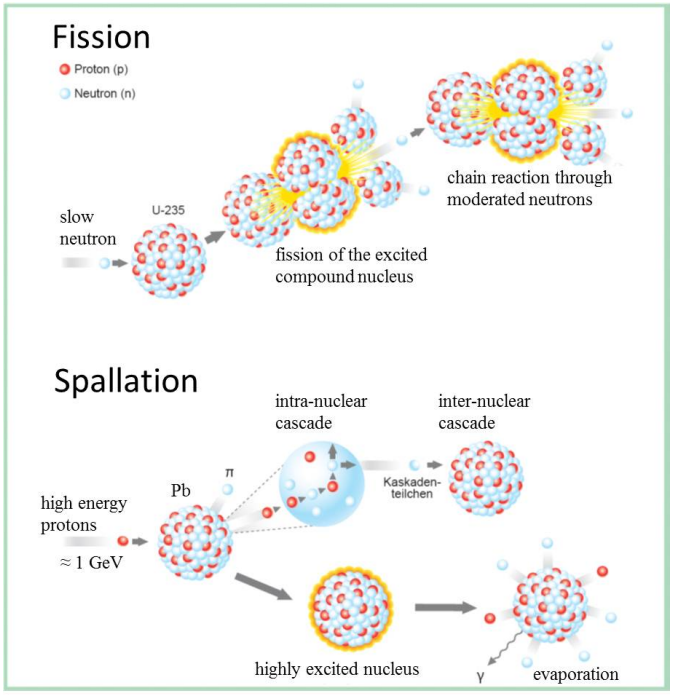


Fig. 8: *A cartoon of the processes of fission and spallation, respectively, used for the production of free neutrons (ess-reports).*

Both processes lead to free neutrons of energies in the MeV region. These neutrons are way too fast to be useful for condensed matter studies. These so-called epithermal neutrons have to be slowed down, which is done most efficiently by collisions with light atoms - e.g. Hydrogen H or Deuterium D in light or heavy water moderators, or C in graphite as in the first reactor, the Chicago Pile 1. During the moderation process after several collisions, the neutrons thermalize and acquire the temperature of the moderator. To adjust the energy spectra to ones need, mainly three types of moderators are being employed:

Moderator	Typical temperature [k]	Neutron energy range [meV]	Neutron wavelength range [Å]
hot source (graphite block)	2500	100 - 1000	0.3 - 0.9
thermal source (H_2O / D_2O)	300	5 - 100	0.9 - 4
cold source (liquid D_2)	25	0.05 - 5	4 - 40

Note that room temperature ~ 300 K corresponds an energy of about 26 meV (1 meV $\hat{=}$ 11.6 K), which is just a typical energy of elementary excitations in a solid. Despite the effort made in these high-tech facilities, the free neutrons available for scattering studies are still extremely rare. In a high flux reactor the neutron flux i. e. the number of neutrons passing through a given area in a given time is in the order of 10^{15} neutrons/cm^2·s. If one compares this value with particle fluxes in gases, the neutron density in high flux sources corresponds to high vacuum conditions of about 10^{-6} mbar pressure. The neutrons have to be transported from the source to the experimental areas, which can either be done by simple flight tubes or so called neutron guides. These are evacuated tubes with glass walls (often covered with metal layers to increase the performance), where neutrons are transported by total reflection from the side, top, and bottom walls in a similar manner like light in glass fibers. The neutron flux downstream at the scattering experiments is then even much lower than in the source itself and amounts to typically 10^6 - 10^8 neutrons/cm^2·s. This means that long counting times have to be taken into account to achieve reasonable statistics in the neutron detector. Just for comparison: the flux of photons of a small Helium-Neon laser with a power of 1 mW (typical for a laser pointer) amounts to some 10^{15} photons/s in a beam area well below 1 mm^2. At modern synchrotron radiation sources, a flux of some 10^{13} photon/s in a similar beamspot can be achieved.

Just as for synchrotron radiation with the XFEL, there is an European project to build the world's most powerful neutron source: the *European Spallation Source ESS*. It will outperform most existing sources by several orders of magnitude in peak flux and allow entirely new experiments to be realized [14]. A presentation of this project will be given by the acting CEO, Colin Carlile, as lecture C9.

3.3 Interaction Processes

The principle probes for condensed matter studies, X-rays, electrons and neutrons feature different interaction processes with matter, leading to a great complementarity. The principle interaction processes are depicted schematically in fig. 9. Details will be given in lecture A4.

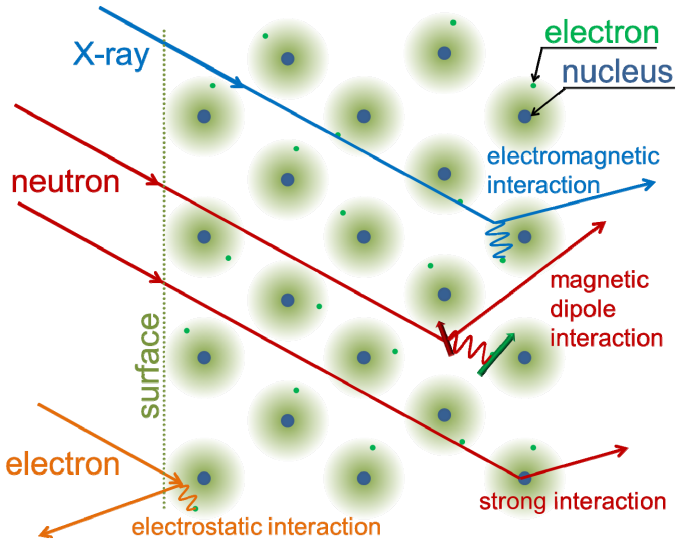


Fig. 9: *Cartoon of the scattering processes of X-rays, neutrons and electrons with atoms in a solid in a 2d representation. The most relevant interaction processes which lead to scattering events are indicated. Note that electrons are mainly scattered in a surface-near region.*

For X-rays, the most relevant scattering process is pure charge or *Thomson scattering* with the differential cross section for scattering from one electron of:

$$\frac{d\sigma}{d\Omega} = r_0^2 \cdot P(\theta) \qquad (21)$$

where $r_0 = \frac{e}{m_0 c^2} = 2.82\ fm$ is the classical electron radius and P(θ) a factor describing the polarization dependence of Hertz' dipole radiation. The Thomson scattering process is the basis for all structural investigations with X-rays since the discovery by Max von Laue one hundred years ago. A single electron is a point-like scatterer, leading to a cross section which is independent of Q, apart from the polarization dependence, Scattering from the extended electron cloud of an atom, on the other hand, leads to a variation of the scattered amplitude with scattering angle described by the so-called *formfactor* - the (normalized) Fourier transform of the electron density of a single atom.

Of course, X-rays as electromagnetic radiation also interact with the spin moment of the electron. This so-called *magnetic x-ray scattering* process is a relativistic correction to charge scattering and typically six orders of magnitude weaker. At absorption edges of elements, the scattering amplitude becomes energy dependent, leading to so-called *anomalous scattering*. Anomalous scattering (see figure 10) enables contrast variation, can also be used to enhance small scattering contributions, such as scattering from charge, orbital or magnetic order, see lecture D11. For most such studies, polarization handling is required, see lecture C7.

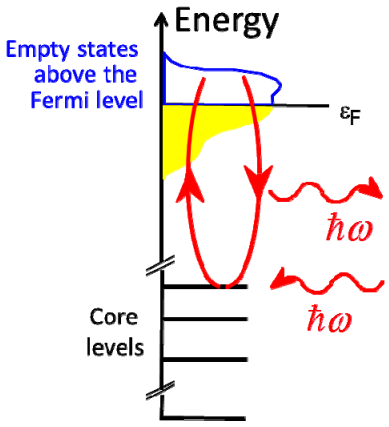


Fig. 10: *Schematic illustration of the second order perturbation process leading to anomalous scattering: core level electros are virtually excited by the incident X-rays into empty states above the Fermi level, if the photon energy is close to an absorption edge. Besides photoelectric absorption, a resonant scattering process can occur, where X-rays of the same wavelength are re-emitted.*

For neutron scattering, two main interaction processes are relevant: scattering with the nucleus due to the strong interaction (*nuclear scattering*) and scattering due to magnetic dipole-dipole interaction between the neutrons magnetic moment and the spin- or orbital moment of unpaired electrons in the solid (*magnetic scattering*).

Since the nucleus is a point-like object compared to the wavelength of thermal neutrons, the differential cross section for nuclear scattering is independent of scattering angle and given by:

$$\frac{d\sigma}{d\Omega} = |b|^2 \tag{22}$$

where b, the scattering length, is a phenomenological parameter as measure of the strength of the interaction potential. b depends not only on the atomic number, but also on the isotope and the nuclear spin orientation relative to the neutron spin.

Magnetic neutron scattering strongly depends on the polarization state of the neutron (for polarization handling, see lecture C6). The differential cross section is given by:

$$\frac{d\sigma}{d\Omega} = \left(\gamma_n r_0\right)^2 \left| -\frac{1}{2\mu_B} \langle \sigma_z' | \underline{\sigma} \cdot \underline{M}_\perp(\underline{Q}) | \sigma_z \rangle \right|^2 \tag{23}$$

The pre-factor $\gamma_n r_0$ has the value $\gamma_n r_0 = 0.539 \cdot 10^{-12}\ cm = 5.39\ fm$. $\underline{\sigma}$ denotes the spin operator, σ_z and σ_z' the polarization state of the neutron before and after the scattering process, respectively. $\underline{M}_\perp(\underline{Q})$ denotes the component of the Fourier transform of the sample magnetization, which is perpendicular to the scattering vector $\underline{Q}$:

$$\underline{M}_{\perp}(\underline{Q}) = \hat{\underline{Q}} \times \underline{M}(\underline{Q}) \times \hat{\underline{Q}}$$
$$\underline{M}(\underline{Q}) = \int \underline{M}(\underline{r}) e^{i\underline{Q}\cdot\underline{r}} d^3 r \qquad (24)$$

This tells us that with neutron scattering we are able to determine the magnetization $\underline{M}(\underline{r})$ in microscopic atomic spatial co-ordinates $\underline{r}$, which allows one not only to determine magnetic structures, but also the magnetization distribution within a single atom.

To obtain an idea of the size of the magnetic scattering contribution relative to nuclear scattering, we can replace the matrix element in (2.3) for a spin ½ particle by the value of one Bohr magneton 1 μ_B. This gives an "equivalent" scattering length for a magnetic scattering of 2.696 fm for a spin ½ particle. This value corresponds quite well to the scattering length of cobalt b_{co} = 2.49 fm, which means that magnetic scattering is comparable in magnitude to nuclear scattering.

3.4 Comparison of Probes

Figure 11 shows a double logarithmic plot of the dispersion relation "wave length versus energy" for the three probes neutrons, electrons and photons. The plot demonstrates, how thermal neutrons of energy 25 meV are ideally suited to determine interatomic distances in the order of 0.1 nm, while the energy of X-rays or electrons for this wavelength is much higher. However, with modern techniques at a synchrotron radiation source, energy resolutions in the meV-region become accessible even for photons of around 10 keV corresponding to a relative energy resolution $\Delta E/E \approx 10^{-7}$ (compare lectures D4, D8 and D9)! The graph also shows that colloids with a typical size of 100 nm are well suited for the investigation with light of energy around 2 eV. These length scales can, however, also be reached with thermal neutron scattering in the small angle region (compare lecture D1). While figure 11 thus demonstrates for which energy-wave-length combination a certain probe is particularly useful, modern experimental techniques extend the range of application by several orders of magnitude.

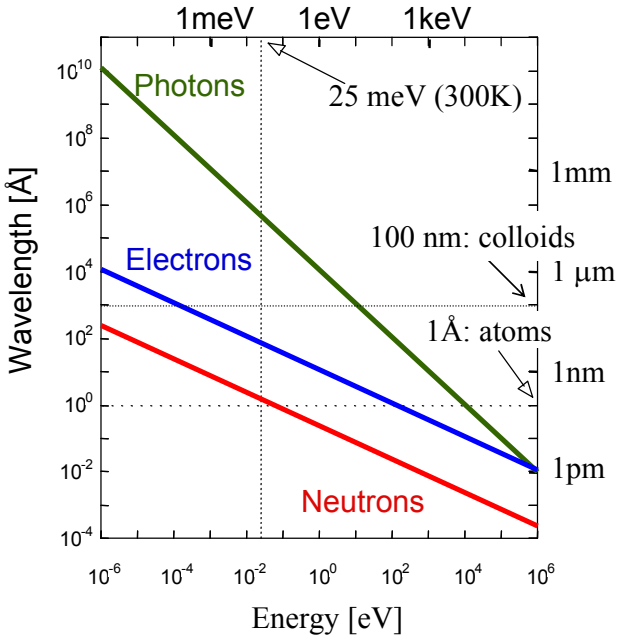


Fig. 11: *Comparison of the three probes - neutrons, electrons and photons - in a double logarithmic energy-wavelength diagram.*

It is therefore useful to compare the scattering cross sections as it is done in figure 12 for X-rays and neutrons. Note that the X-ray scattering cross sections are significantly larger as compared to the neutron scattering cross sections. This means that the signal for x-ray scattering is stronger for the same incident flux and sample size, but that caution has to be applied that the conditions for kinematical scattering are fulfilled. For X-rays, the cross section depends on the number of electrons and thus varies in a monotonic fashion throughout the periodic table. Clearly it will be difficult to determine hydrogen positions with x-rays in the presence of heavy elements such as metal ions. Moreover, there is a very weak contrast between neighbouring elements as can be seen from the transition metals Mn, Fe and Ni in figure 12. However, this contrast can be enhanced by anomalous scattering, if the photon energy is tuned close to the absorption edge of an element (lecture D11). Moreover, anomalous scattering is sensitive to the anisotropy of the local environment of an atom. For neutrons the cross sections depend on the details of the nuclear structure and thus vary in a non-systematic fashion throughout the periodic table. As an example, there is a very high contrast between Mn and Fe. With neutrons, the hydrogen atom is clearly visible even in the presence of such heavy elements as uranium. Moreover there is a strong contrast between the two hydrogen isotopes H and D. This fact can be exploited for soft condensed matter investigations by selectively deuterating certain molecules or functional groups and thus varying the contrast within the sample (see lectures E2, E3).

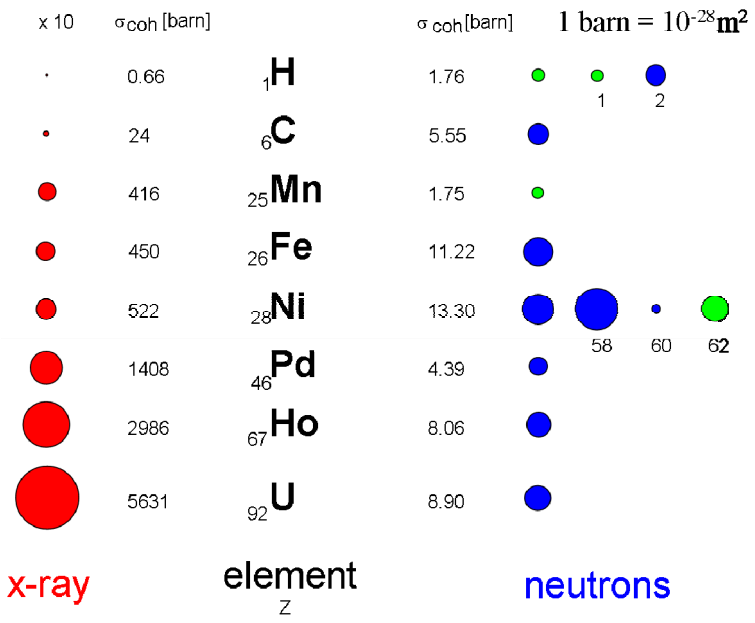


Fig. 12: *Comparison of the coherent scattering cross-sections for x-rays and neutrons for a selection of elements. The area of the coloured circles represent the scattering cross section. In the case of X-rays these areas were scaled down by a factor of 10. For neutrons, the green and blue coloured circles distinguish the cases where the scattering occurs with or without a phase shift of π.*

Finally, both neutrons and X-rays allow the investigation of magnetism on an atomic scale. Magnetic neutron scattering (lectures D3, D4, C6, …) is comparable in strength to nuclear scattering, while non-resonant magnetic X-ray scattering is smaller than charge scattering by several orders of magnitude. Despite the small cross sections, non-resonant magnetic x-ray Bragg scattering from good quality single crystals yields good intensities with the brilliant beams at modern synchrotron radiation sources. While neutrons are scattered from the magnetic induction within the sample, X-rays are scattered differently from spin and orbital momentum and thus allow one to measure both form factors separately. Inelastic magnetic scattering e.g. from magnons or so called quasielastic magnetic scattering from fluctuations in disordered magnetic systems is a clear domain of neutron scattering (lecture D4, D6). Finally, *resonance exchange scattering* XRES, a variant of anomalous X-ray scattering for magnetic systems, allows one not only to get enhanced intensities, but also to study magnetism with element- and band sensitivity (lecture D11).

With appropriate scattering methods, employing neutrons, X-rays or light, processes in condensed matter on very different time and space scales can be investigated. Which scattering method is appropriate for which region within the "scattering vector Q - energy E plane" is plotted schematically in figure 13. Via the Fourier transform, the magnitude of a scattering vector Q corresponds to a certain length scale, an energy to a certain frequency, so that the characteristic lengths and times scales for the various methods can be directly determined from the figure.

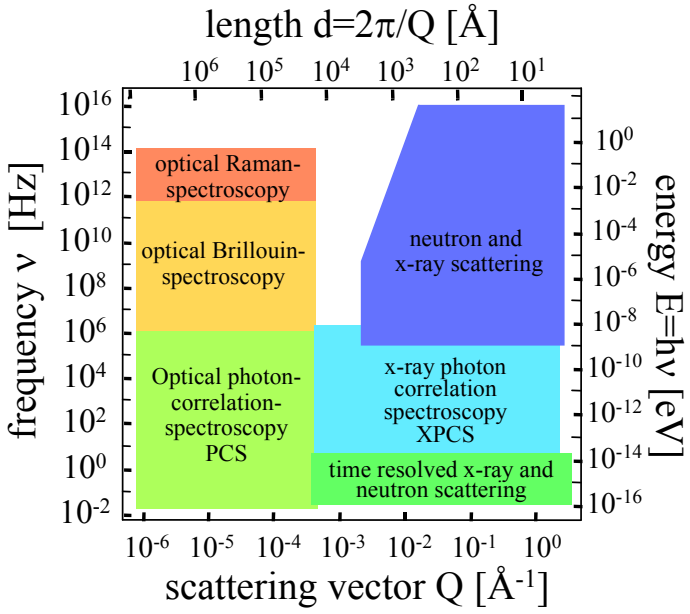


Fig. 13: *Regions in frequency ν and scattering vector Q (ν,Q)- or energy E and length d (E,d)-plane, which can be covered by various scattering methods.*

4 Techniques and Applications

4.1 Introduction

Scattering with electromagnetic radiation (light, soft- and hard-X-rays) and neutrons cover a huge range of energy and momentum transfers (see figure 13), corresponding to an extraordinary range of length- and time scales relevant for research in condensed matter. Exemplary, this is depicted for research with neutrons in figure 14.

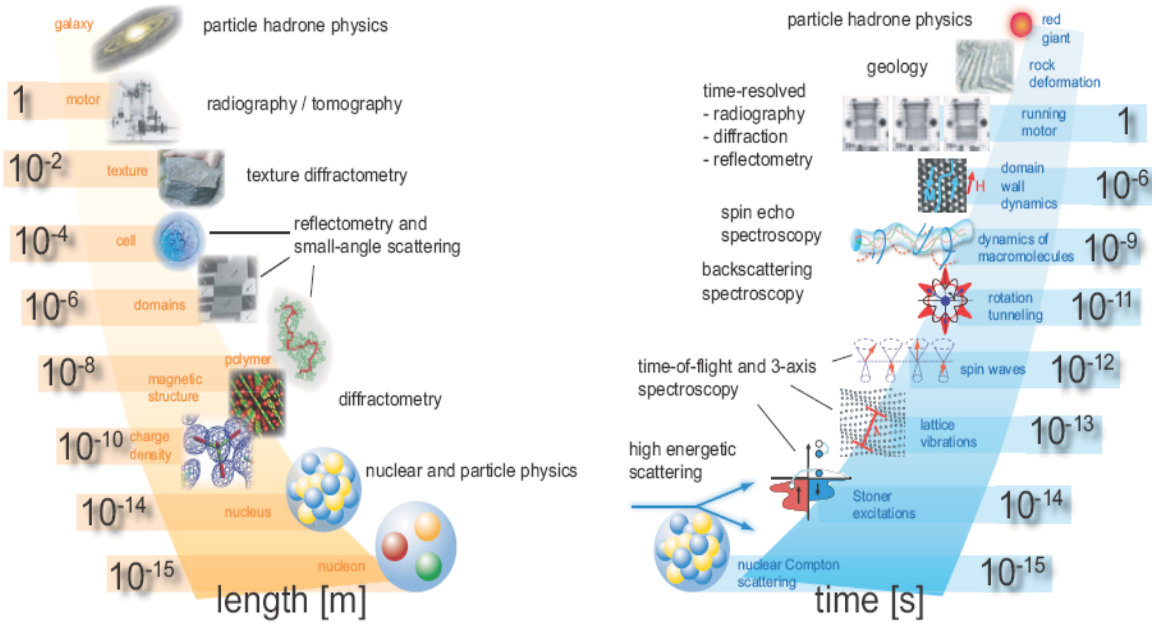


Fig. 14: *Length- and time scales covered by research with neutrons giving examples for applications and neutron techniques.*

The very extremes of length scales - below 10^{-12} m - are the domain of nuclear and particle physics, where e. g. measurements of the charge or the electric dipole moment of the neutron provide stringent tests of the standard model of particle physics without the need of huge and costly accelerators. On the other extreme, neutrons also provide information on length- and time scales relevant for astronomical dimensions, e. g. the decay series of radioactive isotopes produced by neutron bombardment give information on the creation of elements in the early universe. In this course, however, we are only concerned with neutrons as a probe for condensed matter research and therefore restrict ourselves to a discussion of neutron scattering. Still, the various scattering techniques cover an area in phase space from picometers pm up to meters and femtoseconds fs up to hours, a range, which probably no other probe can cover to such an extend.

Different specialized scattering techniques are required to obtain structural information on different length scales:

- With *wide angle diffractometry*, charge (X-rays) or magnetization (neutrons) densities can be determined within single atoms on a length scale of ca. 10 pm[2]. The position of atoms can be determined on a similar length scale, while distances between atoms lie in the 0.1 nm range (lectures B1, D3, D5, D11, E10).
- The sizes of large macromolecules, magnetic domains or biological cells lie in the range of nm to µm or even mm. For such studies of large scale structures, one applies

[2] In this sense, X-rays and neutrons are not only nanometer nm, but even picometer pm probes!

reflectometry (lectures D2, E5) or *small angle scattering technique* (lectures B2, D1, E2) or *imaging* (lectures F1, F2).

- Most materials relevant for engineering or geo-science occur neither in form of single crystals, nor in form of fine powders. Instead they have a grainy structure, often with preferred orientation of the grains. This so called texture determines the macroscopic strength of the material along different directions. *Texture diffractometry* as a specialized technique allows one to determine this granular structure on length scales of up to mm (high energy X-ray diffraction as "3d microscope").
- Finally, for even larger structures, one uses imaging techniques, such as neutron *radiography* or *tomography* (lecture F2), which give a two dimensional projection or full 3-dimensional view into the interior of a sample due to the attenuation of the beam, the phase shift or other contrast mechanisms.

In a similar way, specialized scattering techniques are required to obtain information on the system's dynamics on different time scales:

- *Neutron Compton scattering*, where a high energy neutron in the eV energy range makes a deep inelastic collision with a nucleus in so-called impulse approximation, gives us the momentum distribution of the atoms within the solid. Interaction times are in the femtosecond fs time range.
- With *pump-probe techniques* at free electron lasers, processes in the fs to ps time range can be studied. For this technique one uses the time structure of the radiation and delays a "probe" pulse with respect to the "pump" pulse to study e.g. relaxation processes after excitation in real time.
- In magnetic metals, there exist single particle magnetic excitations, so-called Stoner excitations, which can be observed with inelastic scattering of high energy neutrons using the so-called *time-of-flight spectroscopy* or the *triple axis spectroscopy* technique (lecture D4). Typically, these processes range from fs to several hundred fs.
- The electronic structure of solids, including electronic relaxation processes in the fs time range, can be determined by X-ray spectroscopy techniques (lecture F3, F4).
- Lattice vibrations (phonons) or spin waves in magnetic systems (magnons) have frequencies corresponding to periods in the picosecond ps time range (lecture B4). Again these excitations can be observed with neutron *time-of-flight-*, neutron *triple axis spectroscopy* or at high energy resolution backscattering synchrotron beamlines (lecture D4).
- Slower processes in condensed matter are the tunneling of atoms, for example in molecular crystals or the slow dynamics of macromolecules (lectures B6, E3, E8). Characteristic time scales for these processes lie in the nanosecond ns time range. They can be observed with specialized techniques such as neutron *backscattering spectroscopy*, neutron *spin-echo spectroscopy*, light- or X-ray photon correlation spectroscopy (lecture D8).
- Even slower processes occur in condensed matter on an ever increasing range of lengths scales. One example is the growth of domains in magnetic systems, where domain walls are pinned by impurities. These processes may occur with typical time

constants of microseconds μs. Periodic processes on such time scales can be observed with *stroboscopic scattering* techniques.

- Finally, *time resolved scattering* or *imaging* techniques, where data is taken in consecutive time slots, allow one to observe processes from the millisecond ms to the hour h range.

Even within a spring school of two weeks, it is impossible to cover all scattering techniques and applications. Some will be touched briefly in the application lectures, but we have not foreseen specialized lectures e.g. for texture and strain analysis, or nuclear (neutrons) and electronic (X-rays) Compton scattering.

4.2 Correlation functions

This somewhat advanced section is intended for readers already familiar with scattering and can be skipped during first reading. It is given here for completeness. For details we refer to lecture A5. For sake of simplicity, formulas are given only for neutron scattering, but similar expressions hold for X-ray scattering.

The neutron scattering cross section for nuclear scattering can be expressed in the following form (for simplicity, we restrict ourselves to a mono-atomic system):

$$\frac{\partial^2\sigma}{\partial\Omega\partial\omega} = \frac{k'}{k}\cdot N\cdot\left[\left(\overline{|b|^2}-|\bar{b}|^2\right)S_{inc}(\underline{Q},\omega)+|\bar{b}|^2\ S_{coh}(\underline{Q},\omega)\right] \tag{25}$$

The cross section is proportional to the number N of atoms. It contains a kinematical factor k'/k, i. e. the magnitude of the final wave vector versus the magnitude of the incident wave vector, which results from the phase-space density. The scattering cross section contains two summands: one is the coherent scattering cross section, which depends on the magnitude square of the average scattering length density $|\bar{b}|^2$ and the other one is the incoherent scattering, which depends on the variance of the scattering length $\left(\overline{|b|^2}-|\bar{b}|^2\right)$. The cross section (25) has a very convenient form: it separates the interaction strength between probe (the neutrons) and sample from the properties of the system studied. The latter is given by the so-called scattering functions $S_{coh}(\underline{Q},\omega)$ and $S_{inc}(\underline{Q},\omega)$, which are completely independent of the probe and a pure property of the system under investigation. The *coherent scattering function* $S_{coh}(\underline{Q},\omega)$ (also called *dynamical structure factor* or *scattering law*) is a Fourier transform in space and time of the pair correlation function:

$$S_{coh}(\underline{Q},\omega) = \frac{1}{2\pi\hbar}\int G(\underline{r},t)e^{i(\underline{Q}\cdot\underline{r}-\omega t)}d^3rdt \tag{26}$$

Here the *pair correlation function* $G(\underline{r},t)$ depends on the time dependent positions of the atoms in the sample:

$$G(\underline{r},t) = \frac{1}{N}\sum_{ij}\int \left\langle \delta(\underline{r}' - \underline{\mathrm{r}}_{\mathrm{i}}(0)) \cdot \delta(\underline{r}' + \underline{r} - \underline{\mathrm{r}}_{\mathrm{j}}(t)) \right\rangle d^3 r' \\ = \frac{1}{N}\int \left\langle \rho(\underline{r}',0) \cdot \rho(\underline{r}' + \underline{r},t) \right\rangle d^3 r' \tag{27}$$

$\underline{r}_i(0)$ denotes the position of atom *i* at time *0*, while $\underline{r}_j(t)$ denotes the position of another atom *j* at time *t*. The angle brackets denote the thermodynamic ensemble average, the integral extends over the entire sample volume and the sum runs over all atom pairs in the sample. Instead of correlating the positions of two point-like scatterers at different times, one can rewrite the pair correlation function in terms of the particle density as given in the second line of (27). Coherent scattering arises from the superposition of the amplitudes of waves scattered from one particle at time *0* and a second particle at time *t*, averaged over the entire sample volume and the thermodynamic state of the sample. In contrast, incoherent scattering arises from the superposition of waves scattered from the same particle at different times. Therefore the *incoherent scattering function* $S_{inc}(\underline{Q},\omega)$ is given in the following form:

$$S_{inc}(\underline{Q},\omega) = \frac{1}{2\pi\hbar}\int G_s(\underline{r},t) e^{i(\underline{Q}\cdot\underline{r}-\omega t)} d^3 r dt \tag{28}$$

which is the Fourier transform in space and time of the *self correlation function* $G_S(\underline{r},t)$:

$$G_s(\underline{r},t) = \frac{1}{N}\sum_{j}\int \left\langle \delta(\underline{r}' - \underline{\mathrm{r}}_{\mathrm{j}}(0)) \cdot \delta(\underline{r}' + \underline{r} - \underline{\mathrm{r}}_{\mathrm{j}}(t)) \right\rangle d^3 r' \tag{29}$$

We next define the *intermediate scattering function* $S(\underline{Q},t)$ as the purely spatial Fourier transform of the correlation function (here we have dropped the index “coh” and “inc”, respectively, as the intermediate scattering function can be defined for coherent as well as for incoherent scattering in the same way):

$$S(\underline{Q},t) := \int G(\underline{r},t) e^{i\underline{Q}\cdot\underline{r}} d^3 r \\ = S(\underline{Q},\infty) + S'(\underline{Q},t) \tag{30}$$

For reasons, which will become apparent below, we have separated in the second line the intermediate scattering function for infinite time

$$S(\underline{Q},\infty) = \lim_{t\to\infty} S(\underline{Q},t) \tag{31}$$

from the time development at intermediate times. Given this form of the intermediate scattering function $S(\underline{Q},t)$, we can now calculate the scattering function as the temporal Fourier transform of the intermediate scattering function:

$$S(\underline{Q},\omega)=\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}S(\underline{Q},t)e^{-i\omega t}dt \quad =\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}\left[S(\underline{Q},\infty)+S'(\underline{Q},t)\right]e^{-i\omega t}dt$$
$$=\underbrace{\frac{1}{\hbar}\delta(\omega)S(\underline{Q},\infty)}_{elastic\ scattering}+\underbrace{\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}S'(\underline{Q},t)e^{-i\omega t}dt}_{inelastic\ scattering} \tag{32}$$

In this way, the scattering function has been separated into one term for frequency *0*, i. e. vanishing energy transfer $\Delta E=\hbar\omega=0$ and one term for non-vanishing energy transfer. The first term is the purely elastic scattering, which is given by the correlation function at infinite times. Correlation at infinite times is obtained for particles at rest. A prominent example is the Bragg scattering from a crystalline material, which is purely elastic, while the scattering from liquids is purely inelastic, since the atoms in liquids are moving around freely and thus the correlation function vanishes in the limit of infinite time differences.

Often times the energy of the scattered neutron is not discriminated in the detector. In such experiments, where the detector is set at a given scattering angle, but does not resolve the energies of the scattered neutrons, we measure an *integral cross section* for a fixed direction $\hat{\underline{k}}'$ *of* $\underline{k}'$:

$$\left(\frac{d\sigma}{d\Omega}\right)_{coh,\mathrm{int}}=\int\left.\frac{\partial^2\sigma}{\partial\Omega\partial\omega}\right|_{\hat{\underline{k}}'=const}\cdot d\omega \tag{33}$$

Momentum and energy conservation are expressed by the following kinematic equations of scattering:

$$\hbar\underline{Q}=\hbar\underline{k}'-\hbar\underline{k};\ \hbar\omega=E-E'=\frac{\hbar^2}{2m}\left(k^2-k'^2\right) \tag{34}$$

Due to these kinematic conditions, the scattering vector $\underline{Q}$ will vary with the energy of the scattered neutron E' or the energy transfer $\hbar\omega$ when the integral in (33) is performed. The so-called *quasi-static approximation* neglects this variation and uses the scattering vector $\underline{Q}_0$ for elastic scattering $(\hbar\omega=0)$ in (33). This approximation is valid only if the energy transfer is small compared to the initial energy. This means that the movements of the atoms are negligible during the propagation of the radiation wave front from one atom to the other. In this case, the above integral can be approximated as follows:

$$\left(\frac{d\sigma}{d\Omega}\right)_{coh,QSA}=\frac{k'}{k}\frac{N}{2\pi\hbar}\int\left(\int G(\underline{r},t)e^{i(\underline{Q}_0\cdot\underline{r}-\omega\cdot t)}d^3rdt\right)d\omega$$
$$=\frac{k'}{k}\frac{N}{2\pi\hbar}\int G(\underline{r},t)e^{i\underline{Q}_0\cdot\underline{r}}\delta(t)d^3rdt=\frac{k'}{k}\frac{N}{2\pi\hbar}\int G(\underline{r},0)e^{i\underline{Q}_0\cdot\underline{r}}d^3r \tag{35}$$

which shows that the integral scattering in quasi-static approximation depends on the *instantaneous spatial correlation function* only, i. e. it measures a snapshot of the arrangement of atoms within the sample. This technique is e. g. very important for the determination of short-range order in liquids, where no elastic scattering occurs (see above).

Our discussion on correlation functions can be summarized in a schematic diagrammatic form, see figure 15.

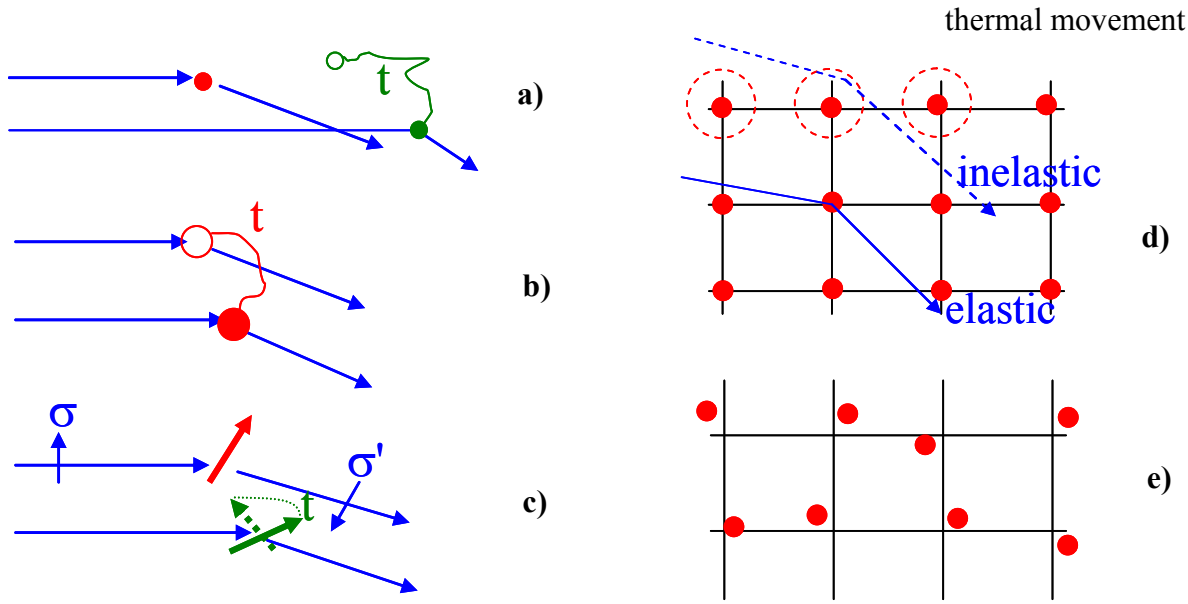


Fig. 15: *Schematic diagrams depicting the various scattering processes: a) coherent scattering is connected with the pair correlation function in space-and time; b) incoherent scattering is connected with the self-correlation function; c) magnetic scattering is connected with the spin pair correlation function; d) elastic and inelastic scattering from a crystal measures average positions and movements of the atoms, respectively, e) inelastic scattering in quasistatic approximation sees a snapshot of the sample.*

Figure 15 shows that coherent scattering is related to the pair correlation between different atoms at different times (15a), while incoherent scattering relates to the one particle self correlation function at different times (15b). In analogy to nuclear scattering, magnetic scattering depends on the correlation function between magnetic moments of the atoms. If the magnetic moment is due to spin only, it measures the *spin pair correlation function* (15c). Since the magnetic moment is a vector quantity, this correlation function strongly depends on the neutron polarization. For this reason, in magnetic scattering we often perform a polarization analysis as discussed in the corresponding lecture C6. Figure 15d depicts elastic and inelastic scattering from atoms on a regular lattice. Elastic scattering depends on the infinite time correlation and thus gives us information on the time averaged structure. Excursions of the atoms from their time averaged positions due to the thermal movement will give rise to inelastic scattering, which allows one e. g. to determine the spectrum of lattice vibrations, see lecture D4 on “inelastic scattering”. Finally, an experiment without energy analysis in quasistatic approximation will give us the instantaneous correlations between the atoms, see figure 15e. This schematic picture shows a snapshot of the atoms on a regular lattice. Their positions differ from the time averaged positions due to thermal movement.

It is evident that specialized techniques are needed to obtain all the information contained in the various cross sections. Optical elements needed to define incident and final wave vectors, and detectors to count the scattered particles, will be discussed in lectures C3, C4, C5, C6 and C7. It is important, however, to realize that these optical elements are never perfect and thus in reality a distribution of initial and final wave vectors are being selected. Therefore, the measured intensity in the detector is not simply proportional to the scattering function $S(\underline{Q},\omega)$ (or more precisely, the cross section), but it is proportional to the convolution of the scattering function (or cross section) with the *experimental resolution function R*:

$$I(\underline{Q}_0,\omega_0) \propto \iint S(\underline{Q},\omega)R(\underline{Q}_0-\underline{Q},\omega_0-\omega)d^3Qd\omega \tag{36}$$

Here, the resolution function *R* appears due to the limited ability of any experimental setup to define an incident or final wave vector $\underline{k}$ or $\underline{k}'$, respectively. *R* therefore depends purely on the instrumental parameters and not on the scattering system under investigation. The art of any scattering experiment is to adjust the instrument - and with it the resolution function - to the problem under investigation. If the resolution of the instrument is too tight, the intensity in the detector becomes too small and counting statistics will limit the precision of the measurement. If, however, the resolution is too relaxed, the intensity will be smeared out and will not allow one to determine the scattering function properly.

4.3 Selected examples for applications

Here we give some selected examples for applications of scattering experiments to topical research, which we selected mainly from our own research at Forschungszentrum Jülich.

4.3.1 Diffractometry

Let us start with structure determination on various length scales. The scattering cross section is related to the Fourier transform of the spatial correlation function and therefore a reciprocal relation exists between characteristic real space distances *d* and the magnitude of the scattering vector *Q*, for which intensity maxima appear:

$$\Delta Q \sim \frac{2\pi}{d} \tag{37}$$

Bragg scattering from crystals provides an example for this equation: the distance between maxima of the Laue function is determined by $\Delta Q \cdot d = 2\pi$, where *d* is the corresponding real space periodicity. (37) is central for the choice of an instrument or experimental set-up, since it tells us which Q-range we have to cover in order to get information on a certain length range in real space.

In order to calculate the corresponding scattering angle, we make use of (3) to calculate 2θ, the scattering angle $\sphericalangle(\underline{k},\underline{k}')$. This angle has to be large enough in order to separate the scattering event clearly from the primary beam. This is why we need different instruments to study materials on different length scales. Table 1 gives two examples.

Example	d	ΔQ	2θ (λ=1 Å)	2θ (λ=10 Å)	Technique
Distance between atoms in crystals	2 Å	3.14 $Å^{-1}$	29°	"cut-off"	wide angle diffraction
Precipitates in metals (e.g. Co in Cu)	400 Å	0.016$Å^{-1}$	0.14°	1.46°	small angle scattering

Tab. 1: *Examples for scattering from structures on different characteristic real space length scales d. ΔQ is the corresponding characteristic scattering vector according to (37), 2θ the scattering angle according to (3), calculated for two different wavelength λ.*

1. The study of structures on atomic length scales is typically done with a wavelength of around 1 Å (comparable to the distance between the atoms) and the scattered intensity is observed at rather large angles between 5° and 175°. Therefore one speaks of *wide angle diffraction*, which is employed for the study of atomic structures.
2. For the study of large scale structures (precipitates, magnetic domains, macromolecules in solution or melt) on length scales of 10 up to 10,000 Å (1 up to 1000 nm), the magnitude of the relevant scattering vectors as well as the corresponding scattering angles are small. Therefore one chooses a longer wavelength in order to expand the diffractogram. The suitable technique is *small angle scattering*, which is employed to study large scale structures.

A topical example for wide angle diffraction comes from the field of superconductivity (lecture E1), which celebrated the 100 anniversary of its discovery just last year, see figure 16.

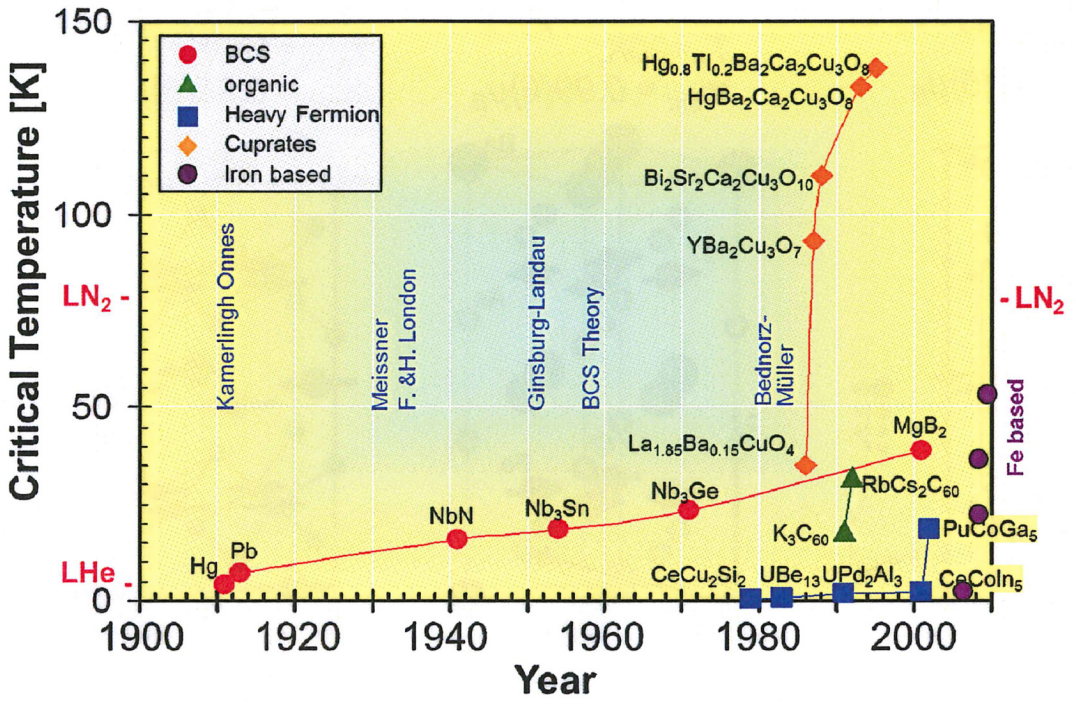


Fig. 16: *A brief history of superconductivity with some examples of superconducting compounds. The critical temperature for the onset of superconductivity T_c is plotted versus the year of discovery.*

The mechanism of superconductivity for the so-called BCS (Bardeen, Cooper, Schrieffer), superconductors is well understood as a Bose-Einstein condensation of Cooper pairs, i.e. electrons bound by the exchange of phonons. More than 25 years ago, Bednorz and Müller (Nobel prize in physics 1987) discovered a new family of superconductors with transition temperatures exceeding liquid nitrogen temperatures, the cuprate superconductors. Despite huge efforts of solid state scientists world wide, the mechanism of superconductivity in these compounds is still unresolved. Therefore it came as a big relieve, when a few years ago a new class of high temperature superconductors was discovered, which is iron based. Just like for cuprates, these superconductors show an intimate proximity to magnetism, i.e. the parent compounds show magnetic order. X-ray and neutron diffraction was employed to determine the structure and magnetic structure of these compounds ([15 - 19]), see figure 17.

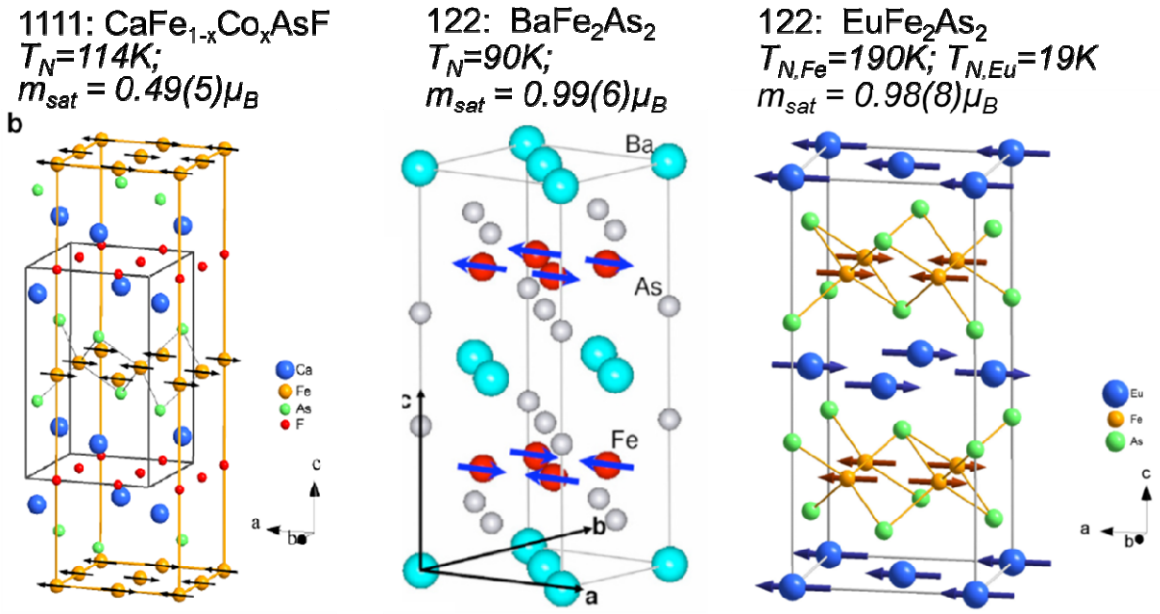


Fig. 17: *Magnetic structures of some iron based superconductors determined by wide angle diffraction .*

This structural information is an essential prerequisite for an understanding of superconductivity in these compounds. All iron-arsenide superconductors have in common that they are layered structures with Fe tetrahedrally surrounded by As. In fact, the closer the bond angles to a perfect tetrahedron, the higher the transition temperature. The magnetic structure is stripe-like with a rather small moment, indicating frustration and/or spin density wave mechanisms.

Wide angle scattering can provide even much more detailed information, such as the magnetization density distribution within the unit cell [20] (for this study, polarized neutrons are required), or the phase diagram as function of temperature, field or pressure [21].

A topical example for small angle scattering stems from the study of magnetic nanoparticles [22]. Magnetic nanoparticles are of fundamental interest for the understanding of magnetism on the nanoscale. They have potential for applications as ferrofluids, in medicine and magnetic data storage. The internal structure of such particles can be determined with a special wide angle scattering technique, the *Pair Distribution Function* PDF analysis (lecture D5). With polarized small angle neutron scattering, the magnetic structure of such particles can be determined, where one finds a significant reduction of the magnetic moment as compared to its bulk value and a spin canting in a surface near layer of the particle [23]. Of special interest is the self assembly of such nanoparticles into 2-dimensional or 3-dimensional ordered structures. We were able to grow well-ordered mesocrystals of maghemite nanoparticles on a silicon wafer substrate. In order to analyse the depth resolved mesocrystal

structure, small angle x-ray scattering is performed under grazing incidence of the beam at a synchrotron radiation facility. Figure 18 shows an example of such a measurement.

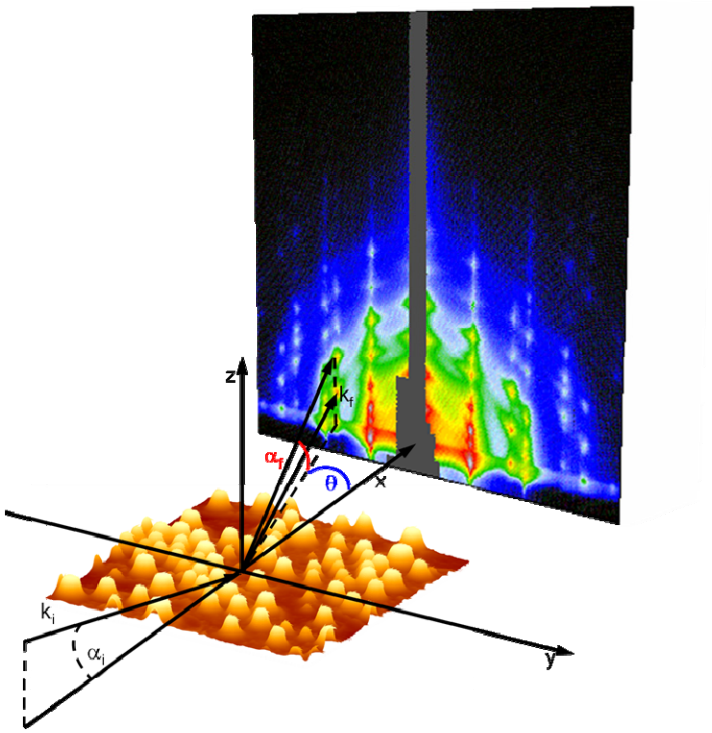


Fig. 18: *Scheme showing the principle of Grazing Incidence Small Angle X-ray Scattering GISAXS with an actual diffraction pattern from mesocrystals of magnetic nanoparticles .*

It turns out that quite complex structures develop. As an example, figure 19 shows the structure obtained for truncated maghemite cubes with 8.5 nm edge length. The body centred tetragonal structure found had been predicted to exist, but had never been observed before.

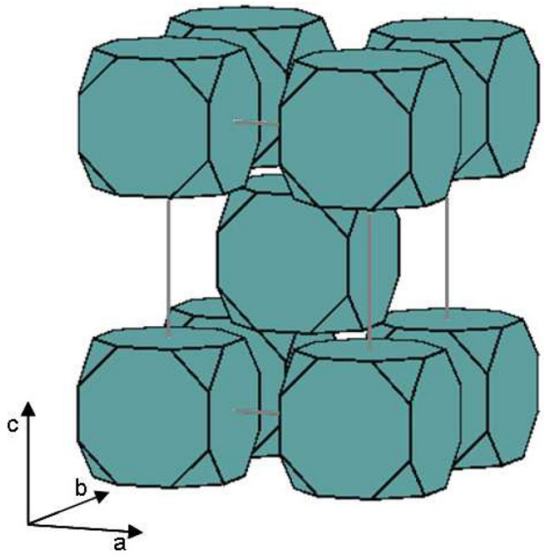


Fig. 19: *Schematic diagram of the body centered tetragonal structure found for truncated nanocubes of maghemite [24].*

From a systematic series of such studies, the interaction energies can be determined. The ability to modulate interparticle interactions by a variation of the particle shape is promising to open a new direction in crystallography, where the mesocrystal structure depends largely on the shape of the primary constituents [24].

4.3.2 Spectroscopy

While diffraction provides information on "where the atoms are", spectroscopy tells us "how the atoms move". Again, this is a very wide field, from diffusion of single atoms or molecules (lecture B3), via coherent elementary excitations in solids (lattice vibrations, spin waves, see lecture B4) and local excitations such as crystal field transitions (lecture B5) to the dynamics of large molecules (lecture B6) like polymers (lecture E3), proteins (lecture E8) or glasses (lecture E9). Just like for diffraction, we can only give two representative examples.

The first example is again concerned with the iron based superconductors. If we want to answer the question, whether the coupling of electrons to Cooper pairs is mediated through phonons, we have to know the phonon dispersion and/or look at charges in the phonon density of states between the superconducting and non-superconducting states. The phonon dispersion is best determined by neutron triple axis or inelastic x-ray spectroscopy (lecture D4). Phonon density of states are measured via time-of-flight neutron scattering (total phonon density of states) or nuclear resonant x-ray scattering (element specific, lecture D9). As an example, figure 20 shows the phonon dispersion of one of the parent compounds of the Fe-As superconductors [25].

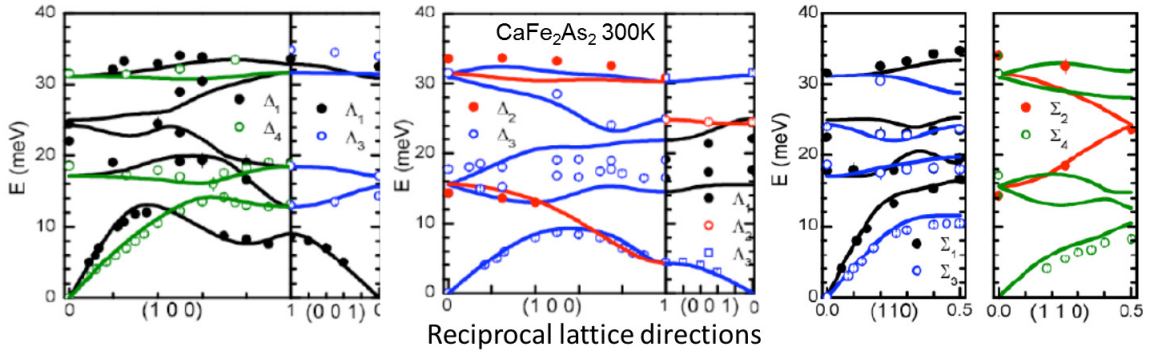


Fig. 20: *Phonon dispersion of $CaFe_2As_2$ at room temperature along main symmetry directions. Data from neutron triple axis spectroscopy are compared to ab-initio calculations (solid lines) [25].*

The calculations largely reproduce the measurements with the exception of some branches, where calculated energies deviate or a strong damping of modes containing only Fe atoms is observed. Similar studies can be performed under applied pressure. Due to the small sample volume, inelastic x-ray scattering is ideally suited to study pressure dependence of phonon modes [26]. Systematic studies of the phonon density of states have been done as well for the non-superconducting as for the superconducting phases [27-30]. At this stage, it is commonly agreed that coupling of electrons to Cooper pairs through phonons alone is not likely to be the mechanism leading to superconductivity in these iron based superconductors. In fact, magnetism has to be involved to reproduce the phonon spectra. Magnetic fluctuations have been observed in the non-superconducting and the superconducting phases and a coupling scheme through magnetic fluctuations, possibly combined with lattice vibrations, is currently being considered.

This example shows clearly, how scattering methods can access microscopic information, which directly relates to the mechanism leading to a macroscopic quantum phenomenon such as superconductivity (see also lecture E1). A similar example can be given from the field of life science for high resolution spectroscopy. Large biomolecules show a slow dynamics in the nsec time range, which requires techniques such as neutron spin echo or photon correlation spectroscopy.

Proteins are the molecular machinery of life. As nanomachines of metabolism, they are in every cell of our body tirelessly active to transport, synthesize, divide and transform substances. The ability of specific proteins to do their job is determined by the sequence of amino acids and their three-dimensional arrangement as determined by x-ray protein crystallography (lecture E10), but also depends on structural rearrangements. To perform their function structural changes are often important. They reach from atomic reorientation to rearrangements of complete domains to enclose substrates, to release products or to reconfigure domains in complexes. Neutron Spin Echo Spectroscopy is a versatile tool to

investigate these large scale movements in biomolecules on different length scales with the ability to determine the timescale of the motions.

The protein alcohol dehydrogenase (ADH) is responsible for the interconversion between alcohol and ketons - a very important catalysis reaction for detoxification after alcohol abuse, see figure 21.

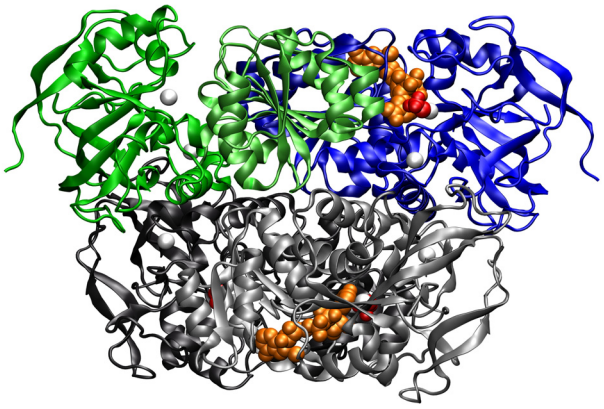


Fig. 21: *The protein alcohol dehydrogenase. The exterior (catalytic) domain tilts outwards and opens the cleft which initiates the catalytic reaction.*

With neutron spin echo spectroscopy, the internal dynamics of the molecule could be studied and the motional amplitude of 0.8 nm determined - an important step in understanding the functionality of proteins in catalytic reactions [31], see lecture E8.

These few selected examples serve to give a taste of the capabilities of scattering methods. More examples will be given in the lectures on topical applications E1 - E9.

With the planned European facilities ESS and XFEL, the future of scattering methods is extremely bright. The high intensity and brightness of these sources will enable entirely new experiments. With neutron chopper spectrometers at ESS, dispersion relations of elementary excitations can be mapped within a few minutes, allowing parametric spectroscopic studies as function of field, pressure or temperature as they are done today for structural studies on diffractometers. Fully coherent scattering of radiation from the XFEL permits the use of oversampling techniques and possibly the determination of the structure of biomolecules without the need for crystallization. The time structure of XFEL with the < 100 fs flashes will allow one to study time dependencies e.g. of chemical reactions in the corresponding time window. The two European projects ESS and XFEL and the possible new types of applications will be presented in lectures C8, C9 and D10.

5 Life at large scale facilities

Neutron and x-ray sources are rather expensive to build and to operate. Therefore, only few such installation exist world wide - especially in the field of research with neutrons, where licensing of nuclear installations is an additional aspect to be considered. Figure 22 shows the geographic distribution of the major facilities for research with neutrons. The Jülich Centre for Neutron Science JCNS is present at some of the world's best sources.

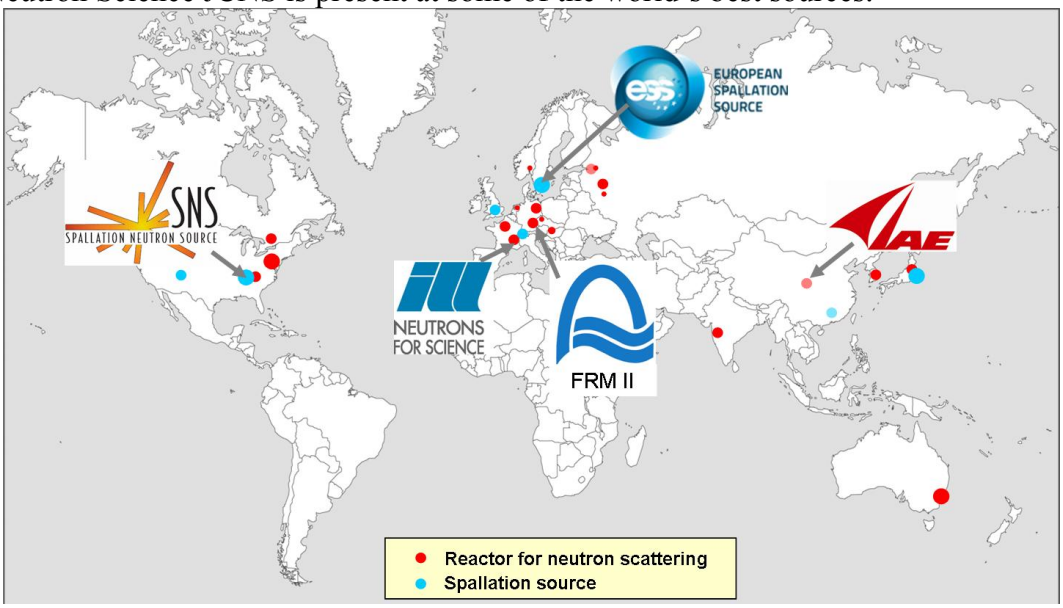


Fig. 22: *Major neutron research centres worldwide which have sources of appreciable flux and a broad instrumentation suite for condensed matter research. JCNS is present at four of the leading sources worldwide: the neutron research reactor FRM II in Garching, Germany, the Institute Laue-Langevin ILL in Grenoble, France, the Spallation Neutron Source SNS in Oak Ridge, USA and the Chinese Advanced Research Reactor CARR close to Beijing, China. JCNS also has a leading involvement in the European Spallation Source project, Lund, Sweden.*

The fact that there are only few sources worldwide implies that experiments at large facilities have to be organized quite different from normal lab-based experiments. Efforts have to be made to use the existing sources as efficient as possible. This means (i) continuous and reliable operation of the source during a large fraction of the year; (ii) many highly performing instruments, which can run in parallel, located around every source; (iii) professional instrument operation with highly qualified staff and a stringent risk management to keep the downtime of instruments and auxiliary equipment as low as possible; (iv) and access for as many scientists as possible. While there are specialized companies which produce beamline and instrument components, there is no true commercial market for neutron or synchrotron instruments. Therefore these instruments are being built by research centres, where usually one or a few staff scientists work closely with engineers and technicians to realize an instrument for a certain application. These highly experienced scientists will then later-on also operate the instruments. The Jülich Centre for Neutron Science JCNS has such

staff scientists located at the outstations at FRM II, ILL and SNS. However, these large scale facilities are way too expensive to be operated just for a small number of scientists. Beamtime is offered to external users from universities, research organizations (such as Max-Planck or Fraunhofer in Germany) and industry. In order for these users to obtain access to a scattering instrument, the user will obtain information from the internet on available instruments, contact the instrument scientist and discuss the planned experiments with the instrument scientist. Once a clear idea and strategy for an experiment has been worked out, the user will write a beamtime proposal where he describes in detail the scientific background, the goal of the planned experiment, the experimental strategy and the prior work. The facility issues a call for proposals in regular intervals, typically twice a year. The proposals received are distributed to members of an independent committee of international experts, which perform a peer review of the proposals and establish a ranking. Typically overload factors between 2 to 3 occur i. e. 2 to 3 times the available beam time is being demanded by external users. Once the best experiments have been selected, the beamtime will be allocated through the facility, where the directors approves the ranking of the committee, the beamline scientist schedules the experiments on the respective instrument and the user office sends out the invitations to the external users. Many facilities will pay travel and lodging for 1 up to 2 users per experiment. It is now up to the user to prepare the experiment as well as possible. If the experiment fails because it was not well prepared, it will be very difficult to get more beamtime for the same scientific problem. Typical experiments last between 1 day and up to 2 weeks. During this time lots of data will be collected which users take home and usually spend several weeks or months to treat the data and model it.

A typical scattering facility will run about 200 days a year with a few hundred visits of user from all over the world. This is also what makes research at large facilities so attractive to young scientists: early-on in their career they will learn to work in large international collaborations, get the opportunity to work on state-of-the-art high-tech equipment and learn to organize their research as efficient as possible. You have therefore chosen well to attend this Spring School!

Conclusion

This overview was meant to give a first introduction to scattering methods, give a climbse of the possibilities provided at current and future sources and outline the structure of the course. You can now look forward to interesting lectures, where many more details will be explained and you will learn the principles to enable you to successfully perform experiments at neutron and synchrotron radiation sources. Have lots of fun and success working with these powerful techniques!

References

[1] W. Friedrich, P. Knipping, M. von Laue, Interferenzerscheinungen bei Röntgenstrahlen, *Sitzungsberichte der Mathematisch-Physikalischen Klasse der Königlich-Bayerischen Akademie der Wissenschaften zu München* (1912) 303 - 322.

[2] W. H. Bragg, W. L. Bragg, The reflection of x-rays by crystals, *Proc. R. Soc. Lond.* **A88** (1913) 428 - 438.

[3] H. Geiger, E. Marsden, On a diffuse reflection of the alpha particles, *Proc. Roy. Soc.* **A82** (1909) 495 - 500.

[4] H. Geiger, E. Marsden, The laws of deflection of the particles through large angles, *Philosophical Magazine Series* **6** (1913) 25604 - 25623.

[5] E. Rutherford, The scattering of alpha and beta particles by matter and the structure of the atom, *Philosophical Magazine* **21** (1911) 669 - 688.

[6] J. Chadwick, Possible existence of a neutron, *Nature* **129** (1932) 312.

[7] J. Chadwick, The existence of a neutron, *Proc. Roy. Soc.* **A136** (1932) 692 - 708.

[8] http://www.atomicarchive.com/History/firstpile

[9] C. G. Shull, Early development of neutron scattering, *Nobel lecture* (1994) www.nobelprize.org.

[10] C. G. Shull, J. S. Smart, Detection of antiferromagnetism by neutron diffraction, *Physical Review* **76** (1949) 1256 - 1257.

[11] B. N. Brockhouse, Slow neutron spectroscopy and a grant enclass of the physical world, *Nobel lecture* (1994). www.nobelprize.org.

[12] R. R. Elder, A. M. Gurewitch et al., Radiation from Electrons in a Synchrotron, *Physical Review* **71** (1947) 829 - 830.

[13] www.XFEL.eu

[14] http://ess-scandinavia.eu/

[15] Y. Su, P. Link, et al., Antiferromagnetic ordering and structural phase transition in $Ba_2Fe_2As_2$ with Sn incorporated from the growth flux, *Physical Review B* **79** (2009) 064504.

[16] Y. Xiao, Y. Su, et al., Magnetic order in the $CaFe_{1-x}Co_xAsF$ (x = 0.00,0.06,0.12) superconducting compounds, *Physical Review B* **79** (2009) 060504.

[17] Y. Xiao, Y. Su, et al., Magnetic structure of $EuFe_2As_2$ determined by single-crystal neutron diffraction, *Physical Review B* **80** (2009) 174424.

[18] J. Herrero-Martín, V. Scagnoli, et al., Magnetic structure of $EuFe_2As_2$ as determined by resonant x-ray scattering, *Physical Review B* **80** (2009) 134411.

[19] S. Nandi, Y. Su, et al., Strong coupling of Sm and Fe magnetism in SmFeAsO as revealed by magnetic x-ray scattering, *Physical Review B* **84** (2011) 054419.

[20] P. J. Brown, T. Chatterji, et al., Magnetization distribution in the tetragonal phase of $BaFe_2As_2$, *Physical Review B* **82** (2010) 024421.

[21] Y. Xiao, Y. Su, et al., Field-induced spin reorientation and giant spin-lattice coupling in $EuFe_2As_2$, *Physical Review B* **81** (2010) 220406.

[22] S. Disch, The spin structure of magnetic nanoparticles and in magnetic nanostructures, *RWTH Aachen University, PhD* (2010).

[23] S. Disch, E. Wetterskog, et al., Quantitative spatial magnetization distribution in iron oxide nanocubes and nanospheres by polarized small angle neutron scattering, *New Journal of Physics* **14** (2012) 013025.
[24] S. Disch, E. Wetterskog, et al., Shape Induced Symmetry in Self-Assembled Mesocrystals of Iron Oxide Nanocubes, *Nano Letters* **11** (2011) 1651-1656.
[25] R. Mittal, L. Pintschovius, et al., Measurement of Anomalous Phonon Dispersion of $CaFe_2As_2$ Single Crystals Using Inelastic Neutron Scattering, *Physical Review Letters* **102** (2009) 217001.
[26] R. Mittal, R. Heid, et al., Pressure dependence of phonon modes across the tetragonal to collapsed-tetragonal phase transition in $CaFe_2As_2$, *Physical Review B* **81** (2010) 144502.
[27] R. Mittal, Y. Su, et al., Inelastic neutron scattering and lattice-dynamical calculations of $BaFe_2As_2$, *Physical Review B* **78** (2008) 104514.
[28] R. Mittal, S. Rols, et al., Phonon spectra in $CaFe_2As_2$ and $Ca_{0.6}Na_{0.4}Fe_2As_2$: Measurement of the pressure and temperature dependence and comparison with ab initio and shell model calculations, *Physical Review B* **79** (2009) 144516.
[29] R. Mittal, M. Zbiri, et al., Effects of magnetic doping and temperature dependence of phonon dynamics in $CaFe_{1-x}Co_xAsF$ compounds (x = 0, 0.06, and 0.12), *Physical Review B* **79** (2009) 214514.
[30] M. Zbiri, R. Mittal, et al., Magnetic lattice dynamics of the oxygen-free FeAs pnictides: how sensitive are phonons to magnetic ordering?, *Journal of Physics-Condensed Matter* **22** (2010) 8.
[31] R. Biehl, B. Hoffmann, et al., Direct Observation of Correlated Interdomain Motion in Alcohol Dehydrogenase, *Physical Review Letters* **101** (2008) 138102.

A 2 Scattering Theory: Born Series [1]

Stefan Blügel
Peter Grünberg Institut and
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Since Rutherford's surprise at finding that atoms have their mass and positive charge concentrated in almost point-like nuclei, scattering methods are of extreme importance for studying the properties of condensed matter at the atomic scale. Electromagnetic waves and particle radiation are used as microscopic probes to study a rich variety of structural and dynamical properties of solids and liquids. Atomistic processes in condensed matter take place at length scales on the order of an Ångström (1Å$= 10^{-10}$m) and an energy scale between a meV and a few eV. Obviously, detailed information concerning atomic systems require measurements related to their behavior at very small separations. Such measurements are in general not possible unless the de Broglie wavelength ($\lambda = \frac{h}{p} = \frac{h}{mv}$) of the relative motion of the probing particle is comparable to these distances. This makes x-ray scattering and neutron scattering, in addition to electron scattering and to a certain extent also Helium scattering, to the outstanding microscopic "measurement instruments" for studying condensed matter. To push electromagnetic waves in this area one uses either x-rays with wavelengths of a few Ångströms, but in the keV energy range, or light with energies in the eV range, but wavelengths of some 1000 Å. Neutrons (and Helium atoms) make it possible to match energy and wavelengths simultaneously to the typical atomic spacings and excitation energies of solids (solid surfaces). Thus, a simultaneous spatial and temporal resolution of atomistic or magnetic processes is possible. In addition, the photon, neutron, electron and under certain conditions also Helium possess internal degrees of freedom such as a polarization vector or a spin with which the probes couple to core and valence electrons. The photon, neutron and Helium result in only week interaction with matter, which simplifies considerably the analysis and interpretation of experiments as multiple scattering processes are frequently of minor importance and can often be ignored completely, which makes an interpretation of the scattering results valid within a kinematic scattering theory.

In this Chapter, we will provide a brief introduction to the elementary concepts and methodology of scattering theory. The focus lies on the introduction of the description of the scattering process in terms of the Hamiltonian of the scattering projectile at a finite range interaction potential of a single site target by the method of partial waves, i.e. using differential equation methods and the Lippmann-Schwinger equation, i.e. using an integral equation formulation of scattering that leads then to the first Born approximation of scattering and the distored wave Born approximation. The former is the approximation of choice if multiple scattering is unimportant and the latter is applied in the analysis of grazing-incidence small-angle scattering experiments discussed in more details chapter D2. The lecture closes with the discussion of the scattering on the lattice rather than a single site target resulting briefly the Bragg scattering, that will be discussed in more detail in the lecture Scattering Theory: Dynamical Theory (A3). The subject is typically part of an Advanced Quantum Mechanics curriculum and is therefore elaborated at textbooks on quantum mechanics. A selection is given as references [1, 2, 3, 4].

2 The Scattering Problem

In a scattering experiment a beam of particles is allowed to strike a target,[2] and the particles that emerge from the target area or *scattering volume*, respectively, are observed.

[2] In the language of elementary scattering theory one frequently refers to a target although we keep in mind that in the language of condensed matter, it is referred to as the sample.

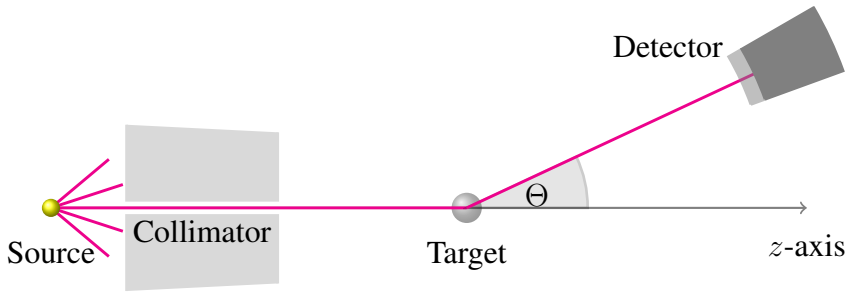


Fig. 1: *A set-up of a typical diffraction experiment consisting of a particle source, a scattering target and a particle detector. The beam passes through a collimator with an beam opening* Δr.

2.1 The Experimental Situation

A schematic representation of a standard scattering experiment appears in Fig. 1. Each scattering experiment consists of three indispensable elements: (i) The source of incident beam of particles or electromagnetic wave, to propagate with wave vector $\mathbf{k} = \frac{2\pi}{\lambda}\hat{\mathbf{k}}$ of wavelength λ along the direction $\hat{\mathbf{k}}$, which we assume without loss of generally to be the $\hat{\mathbf{z}}$-direction (the axis of the collimator). (ii) The target, that we consider stationary, a reasonable assumption in condensed matter physics and (iii) the detector, whose function is to simply count the number of particles of a particular type that arrive at its position $\mathbf{r}$ with the coordinate $\mathbf{r} = (x, y, z)$ in real space along the direction $\hat{\mathbf{r}}$ at the angle (θ, φ) with respect to the axis of the propagating incident beam. Ideally it may be set to count only particles of a given energy, spin or polarization vector, respectively. We assume throughout that the source and detector are classical objects that have a clearly defined, precisely controllable effect on the scattering process. The detector will be assumed to be 100% efficient and to have no effect on the scattered particle prior to the time it enters the detector.

The result of the scattering experiment will vary with the energy E of the incident beam. In order to simplify the analysis of the experimental data, the energy spectrum of the incident beam should be sharply peaked so that the experiment may be considered to take place at a unique energy eigenvalue E. To this end, in most experiments care is taken to achieve a monochromatic incident beam characterized by the wave vector $\mathbf{k}$. We shall assume here that the beam emerging from the collimator is both perfectly monochromatic and perfectly collimated, as well. Of course, according to the uncertainty principle, a beam of finite cross section (of the size of the collimator opening Δr) cannot be perfectly monochromatic ($\Delta k > 0$) and perfectly collimated as well. We may, however, assume the beam to be sufficiently well collimated that the angular divergence may be ignored in an actual experiment. In that case we must necessarily have not a monochromatic beam represented by a plane wave, but rather a beam describable as a superposition of such waves. In this respect the collimator can be considered the fourth indispensable element of a scattering experiment. It shields the detector from the incoming beam to the ideal extent that no count is measured in the detector without target, and produces a small beam of monochromatic energy. We further assume that the detector has a small opening angle $d\Omega$ and is positioned at large distance from the target. Under these conditions, the scattered beam can be characterized at the position of the detector by the wave vector $\mathbf{k}'$ and energy E'. Summarizing, in a scattering experiment a wave packet of incident particles characterized by the initial state $(\mathbf{k}, E, \mathbf{e})$, denoting the polarization vector $\mathbf{e}$ of an x-ray beam as a representative of an internal

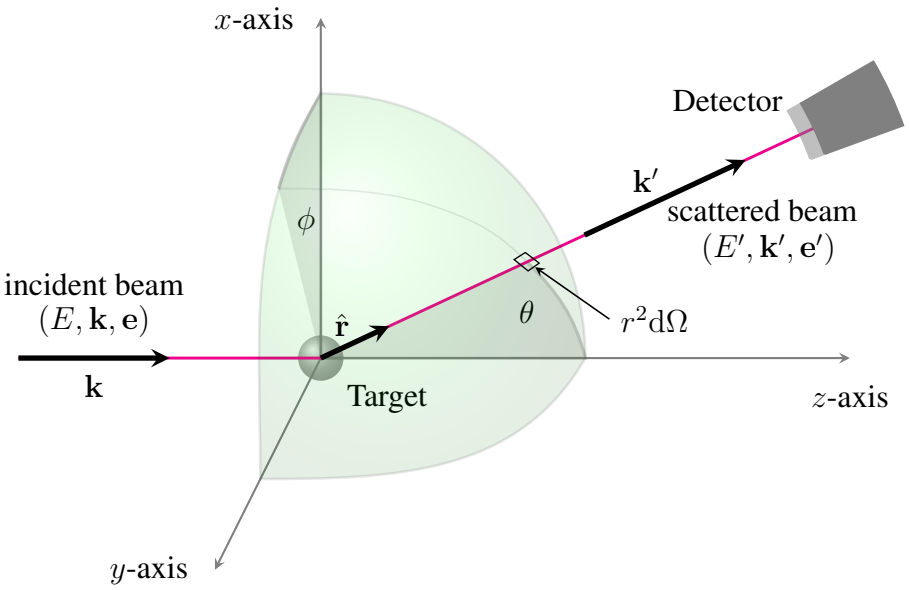


Fig. 2: *The geometry of the scattering experiment.*

degree of freedom of the particle, is scattered into the final state:

$$(\mathbf{k}, E, \mathbf{e}) \overset{\text{scattering}}{\longrightarrow} (\mathbf{k}', E', \mathbf{e}'). \tag{1}$$

The scattering process is characterized by the *scattering vector*

$$\mathbf{Q} = \mathbf{k}' - \mathbf{k} \tag{2}$$

and the *energy transition*

$$\hbar\omega = E' - E. \tag{3}$$

$\hbar\mathbf{Q}$ represents the momentum transfer during scattering, since according to de Broglie, the momentum of the particle corresponding to the wave with wave vector $\mathbf{k}$ is given by $\mathbf{p} = \hbar\mathbf{k}$. For *elastic scattering* (*diffraction*), it holds that $E' = E$ and $|\mathbf{k}'| = |\mathbf{k}|$ and all possible scattering vectors are located on a sphere, called *Ewald-sphere*. Structural investigations are always carried out by elastic scattering. The magnitude Q of the scattering vector can be calculated from wavelength λ and scattering angle θ as follows

$$Q = \sqrt{k^2 + k'^2 - 2kk'\cos\theta} = \sqrt{2k^2(1-\cos\theta)} = k\sqrt{2(1-\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2})} = \frac{4\pi}{\lambda}\sin\frac{\theta}{2} \tag{4}$$

2.2 Description of Scattering Experiment

After the beam of particles is emitted from the collimated source, the experimenter has no control over the particles until they have reached his detector. During that time, the propagation is controlled solely by the laws of quantum mechanics and the Hamiltonian of the projectile-target system. We restrain our description to the nonrelativistic domain and may thus formulate the physical situation in terms of the solution of a Schrödinger equation using an appropriate

Hamiltonian and suitable boundary conditions. The occurring phenomena can be very complex. We assume in the following that the incident particles do not interact with each other during the time of flight, rather that the particles fly one by one, and that incident particles and the target particles do not change their internal structures or states, but only scatter off each other. Internal excitations, rearrangements, charge or spin exchange are also excluded. In fact, internal degrees of freedom such as the spin or polarization vector are currently completely neglected. That means we shall consider only the purely *elastic scattering*. We further neglect the *multiple scattering* in the target and consider at first only the interaction of an incident particle with one target particle, the state of both is described by a two-particle wave function $\Psi(\mathbf{r}_\mathrm{P}, \mathbf{r}_\mathrm{T})$ and the interaction is described by a potential $V(\mathbf{r})$ which depends only on the relative distance $\mathbf{r} = \mathbf{r}_\mathrm{T} - \mathbf{r}_\mathrm{P}$, where the subscripts T and P denote the target and the incoming particle, respectively. We shall confine ourselves in this article to scattering processes in which only short-range central forces are present. In the presence of such potentials, the particle is not under influence of the target potential when they are emitted from the source or when they enter the detector. Since for the two-body problem the motion of the center of mass $\mathbf{R}$ can be separated out, the problem reduces to the scattering of a particle with the reduced mass $\frac{1}{m} = \frac{1}{m_\mathrm{T}} + \frac{1}{m_\mathrm{P}}$ at the potential $V(\mathbf{r})$. Since the potential does not depend explictly on the coordinate of the center of mass $\mathbf{R}$, the two-particle wave function can be expressed in terms of a product of single-particle wave functions $\Psi(\mathbf{r}_\mathrm{P}, \mathbf{r}_\mathrm{T}) = \phi(\mathbf{R})\psi(\mathbf{r})$,[3] both solutions of two separate Schrödinger equations. The elementary two-body scattering process could be intrinsically elastic, but recoil of the target particle might lead to a transfer of energy to the target. In the present context elastic scattering specifically excludes such effects. Considering a solid as target, depending on the energy of the projectile and the interaction of the constituent atoms in a solid, this can be a very good assumption, as for favorable circumstances all atoms contribute to the scattering mass of the solid (of course, in other chapters it becomes clear that atoms in a solid vibrate and a scattering event may cause inelastic excitations of phonons in the vicinity of the elastic energy). To think that the target particle or a target solid is infinitely heavy relative to the mass of the incident particle simplifies our thinking further. In this case the center of mass of the target remains stationary at the position of the target particle throughout the scattering process. Under these circumstances the relative coordinate $\mathbf{r}$ represents the actual laboratory coordinate of the light particle and the mass m is then the mass of the incident particle. These parameters enter the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi \qquad \text{with} \qquad V(\mathbf{r}) = 0 \qquad \text{except } \mathbf{r} \in \text{ target region } \mathbb{T} \quad (5)$$

to be solved.

The time-dependent Schrödinger equation seems a natural starting point for the description of a scattering event as it is not a stationary process but involves individual discrete particles as projectiles, but as we see a bit later the good news is, that under reasonable assumptions that are fulfilled in typical experimental situations, the same results are obtained using a time-independent description applying the stationary Schrödinger equation.

At the vicinity of the collimator and detector, the solution of the potential-free Schrödinger equation (5) is analytically known as the free-particle wave packet:

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi)^3}\int d^3k\, A(\mathbf{k})\,\psi_\mathbf{k}(\mathbf{r}, t) \qquad \text{with} \qquad \psi_\mathbf{k}(\mathbf{r}, t) = e^{i\mathbf{kr}}e^{-i\frac{\hbar k^2}{2m}t}. \quad (6)$$

[3] This does not hold if target and projectile are identical particles. Identical articles scattering about angle θ and $\pi - \theta$ cannot be distinguished.

The wave packet is expressed as a superposition of a complete set of stationary-state solutions $\psi_{\mathbf{k}}(\mathbf{r})$ of this Schrödinger equation, which are plane waves $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}$. The energy eigenvalue corresponding to the eigenfunction $\psi_{\mathbf{k}}$ is simply $E_{\mathbf{k}} = \frac{\hbar^2}{2m}k^2$. For convenience the target is placed at the origin of the coordinate system ($r = 0$). The coefficient $A(\mathbf{k})$ is the probability amplitude for finding the wave number $\mathbf{k}$, or momentum $\hbar\mathbf{k}$ in the initial state. We assume that the properties of the source are such that the wave packet is close to monochromatic and that the amplitude function $A(\mathbf{k})$ peaks about the average momentum $\hbar\mathbf{k} = \hbar\mathbf{k}_\circ$ with a spread in the wave number $\Delta\mathbf{k}$ that is small compared to $\mathbf{k}_\circ$ ($\Delta\mathbf{k} \ll \mathbf{k}_\circ$), related to the opening of the collimator $\Delta\mathbf{r}$ by Heisenberg's uncertainty principle ($\Delta\mathbf{k}\Delta\mathbf{r} \simeq 1$). If we finally impose the condition $\frac{(\Delta\mathbf{k})^2}{\mathbf{k}_\circ}L \ll 1$ (equivalent to the condition $\frac{\lambda_\circ}{\Delta r} \ll \frac{\Delta r}{L}$) the packet does not spread appreciably during the course of the experiment with the set-up of length L, and we can finally show that a stationary description of the scattering problem is sufficient. Under these conditions the energy-dependent phase factor of $\psi_{\mathbf{k}}(\mathbf{r}, t)$ can be conveniently be approximated about the median energy $k^2 \simeq -k_\circ^2 + 2\mathbf{k}\cdot\mathbf{k}_\circ$ and Eq. (6) will become

$$\psi(\mathbf{r}, t) = e^{i\frac{\hbar k_\circ^2}{2m}t}\,\frac{1}{(2\pi)^3}\int d^3k\, A(\mathbf{k})\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{v}_\circ t)} = e^{i\frac{\hbar k_\circ^2}{2m}t}\psi(\mathbf{r} - \mathbf{v}_\circ t, 0), \tag{7}$$

i.e. a wave packet $\psi(\mathbf{r}, t)$ centered about the origin $r = 0$ at $t = 0$ moves at the classical velocity $\mathbf{v}_\circ = \frac{\hbar\mathbf{k}_\circ}{m}$ and the packet at time $t > 0$ will have exactly the same shape, but centered about $\mathbf{r} = \mathbf{v}_\circ t$. Thus, the initial state $\lim_{t\to-\infty}\psi(\mathbf{r}, t)$ and the final state $\lim_{t\to+\infty}\psi(\mathbf{r}, t)$ can be expressed by Eq. (7), but with the coefficients $A(\mathbf{k})$ in the final state having been modified compared to the ones in the initial state due to the scattering, as we shall discuss immediately.
A solution of Eq. (5) requires, however, the specification of boundary conditions imposed on the solution that reflect the physical situation in the laboratory as discussed in section 2.1. The proper boundary condition is a condition on the wave function when the particle and target are far apart. It can be motivated from *Huygens' principle* [5] who proposed that every point which a luminous disturbance reaches becomes a source of a spherical wave, and the sum of these secondary waves determines the form of the wave at any subsequent time. For a single target scatterer we express the wave function

$$\psi_{\mathbf{k}}(\mathbf{r}) \stackrel{r\to\infty}{\longrightarrow} e^{i\mathbf{kr}} + \frac{1}{r}e^{ikr}f_{\mathbf{k}}(\theta, \phi) \qquad \forall\mathbf{k} \quad \text{and} \quad t > 0 \tag{8}$$

in terms of a superposition for the incoming wave plus an outgoing scattered wave emanating from the target, removing some of the incoming particles from the incident primary beam. $f(\theta, \phi)$, $f(\hat{\mathbf{r}})$ or $f(\hat{\mathbf{k}}')$, respectively, denotes the scattering amplitude. This form of the wave function is motivated by the fact that we expect, after scattering, an outgoing spherical wave, modified by the scattering amplitude, interfering with the incoming wave; we will later show a more rigorous justification of this expression. Consistent to the lab schematics in Fig. 1, $\mathbf{k}_\circ \| \hat{\mathbf{z}}$ and the azimuthal and polar scattering angle (θ, ϕ) are given by the projection of the direction of the wavevector $\hat{\mathbf{k}}$ of the scattered wave, e.g. into the detector at direction $\hat{\mathbf{r}}$ and the $\hat{\mathbf{z}}$ direction. If we replace $\psi_{\mathbf{k}}(\mathbf{r}, t)$ in (6) by its asymptotic form given in (8), when the packet is far from the target, the wave function $\psi'(\mathbf{r}, t)$ (where we use a prime to denote the wave function after scattering) breaks up into two terms

$$\psi'(\mathbf{r}, t) = \psi_{\mathbf{k}}(\mathbf{r}, t) + \psi_{\rm sc}(\mathbf{r}, t), \qquad \text{with} \quad \psi_{\rm sc}(\mathbf{r}, t) = 0 \quad \text{for} \quad t < 0 \tag{9}$$

with the incident beam $\psi_{\mathbf{k}}(\mathbf{r}, t)$ identical to Eq. (7) and the scattered wave $\psi_{\rm sc}(\mathbf{r}, t)$ according

to Eq. (6)

$$\psi_{\rm sc}(\mathbf{r},t) = \frac{1}{r}\,\frac{1}{(2\pi)^3}\int d^3k\, f_{\mathbf{k}}(\hat{\mathbf{r}})A(\mathbf{k})\,e^{ikr}\,e^{-i\frac{\hbar k^2}{2m}t}. \tag{10}$$

If we assume now that the scattering amplitude is slowly varying over the spread of the wave numbers $\Delta\mathbf{k}$, and thus approximate $f_{\mathbf{k}}(\hat{\mathbf{r}}) \simeq f_{\mathbf{k}_\circ}(\hat{\mathbf{r}})$ as well as making use of the approximations

$$k \simeq \hat{\mathbf{k}}_\circ \cdot \mathbf{k} \qquad \text{and} \qquad k^2 \simeq -k_\circ^2 + 2\mathbf{k}\cdot\mathbf{k}_\circ \qquad \text{for} \qquad (\Delta k)^2 \ll k_\circ^2 \tag{11}$$

($k = |\mathbf{k}| = |\mathbf{k}_\circ + \Delta\mathbf{k}| = [k_\circ^2 + 2\mathbf{k}_\circ\cdot\Delta\mathbf{k} + (\Delta k)^2]^{1/2} \simeq k_\circ[1 + 2\mathbf{k}_\circ\cdot\Delta\mathbf{k}/k_\circ^2]^{1/2} \simeq k_\circ + \hat{\mathbf{k}}_\circ\cdot(\mathbf{k}-\mathbf{k}_\circ)$) in which k is equal to the projection along $\mathbf{k}_\circ$ we obtain

$$\psi_{\rm sc}(\mathbf{r},t) = e^{-i\frac{\hbar k_\circ^2}{2m}t}\,\frac{1}{r}\,f_{\mathbf{k}_\circ}(\hat{\mathbf{r}})\,\frac{1}{(2\pi)^3}\int d^3k\,A(\mathbf{k})\,e^{i\mathbf{k}\cdot(\hat{\mathbf{k}}_\circ r - \mathbf{v}_\circ t)} = e^{-i\frac{\hbar k_\circ^2}{2m}t}\,\frac{1}{r}\,f_{\mathbf{k}_\circ}(\hat{\mathbf{r}})\,\psi(\hat{\mathbf{k}}_\circ r - \mathbf{v}_\circ t, 0)\,. \tag{12}$$

Thus, after the incident packet has passed the target a spherical scattered wave shell of thickness Δr, equal to the size of the packet, centered on the origin and having a radius $r = v_\circ t$ emerges from the target. One finds further that the incoming wave packet given in (7) and the scattered wave, Eq. (12), share absolutely the same time dependence and are the same for all k. The solution is actually a superposition of all available wave numbers according to $\int d^3k A(\mathbf{k})\ldots$. Since there is no mode-mode coupling such as $\mathbf{k} \rightarrow \mathbf{k_1} + \mathbf{k_2}$, it is totally sufficient to solve the problem in terms of a scattering problem of $\psi_{\mathbf{k}}(\mathbf{r})$ on the basis of a stationary Schrödinger equation for all relevant wave vectors $\mathbf{k}$, which will be pursued during the rest of the manuscript, and keep thereby in mind that a wave packet is formed with a certain probability amplitude. This stationary problem with plane waves as incident beam simplifies the description of scattering significantly.

2.3 Coherence

The formation of a wave packet bears, however, a consequence on which we shall briefly touch upon: The scattering pattern or diffraction pattern, respectively, will be a superposition of patterns for different incident wave vectors $(\mathbf{k}, \mathbf{k}+\Delta\mathbf{k})$ and the question arises, which information is lost due to these non-ideal conditions. This "instrumental resolution" is intimately connected with the "coherence" of the beam and the size of the *scattering volume* in comparison to the *target volume*. Coherence is needed, so that the interference pattern is not significantly destroyed. Coherence requires a phase relation between the different components of the beam. A measure for the coherence length l is given by the distance, at which two components of the beam become fully out of phase, i.e. when one wave train at position $\mathbf{r}$ exhibits a maximum, meats a wave train exhibiting a minimum, thus experiencing a phase difference of $\lambda/2$. If the coherence length $l_\parallel$ is determined by the wavelength spread, λ and $\lambda + \Delta\lambda$ one refers to the temporal or longitudinal coherence. The condition $l_\parallel = n\lambda = (n - \frac{1}{2})(\lambda + \Delta\lambda)$ translates then into

$$l_\parallel = \frac{1}{2}\frac{\lambda^2}{\Delta\lambda} \qquad \text{for longitudinal coherence}$$

and

$$l_\perp = \frac{1}{2}\frac{\lambda}{\Delta\theta} \qquad \text{for transversal coherence.}$$

Analogously one obtains the transversal coherence length $l_\perp$ shown in above equation due to the divergence of the beam $\Delta\theta$ that results from the finite transverse beam size due to the finite extension of the source. In many instruments, the vertical and horizontal collimations are different and the vertical one can even be different along different spatial directions.
Together, the longitudinal and the two transversal coherence lengths define a coherence volume. This is a measure for a volume within the sample, in which the amplitudes of all scattered waves superimpose to produce an interference pattern. Normally, the coherence volume is significantly smaller than the sample size, typically a few 100 Å for neutron scattering, up to μm for synchrotron radiation. Scattering between different coherence volumes within the sample is no longer coherent, i.e. instead of the amplitudes, the intensities of the contributions to the scattering pattern have to be added. This limits the spatial resolution of a scattering experiment to the extension of the coherence volume.

2.4 The Cross Section

A general measure of the scattered intensity $I(\Omega)$ is the *differential cross section* $(\frac{d\sigma}{d\Omega})$. It is defined by the number of particles dN counted per unit time dt scattered into a cone of solid angle $d\Omega = \sin\theta\, d\theta\, d\phi$ in the detector located at the distance r along a ray specified by the direction $\hat{\mathbf{r}}$, at angle (θ, ϕ) or the solid angle Ω, respectively, normalized to the current of the incoming particles $j_{\rm in}$

$$\frac{1}{j_{\rm in}}\frac{dN}{dt} = \left(\frac{d\sigma}{d\Omega}\right) d\Omega = \left(\frac{d\sigma}{d\Omega}\right)\frac{1}{r^2} dA. \tag{13}$$

$d\sigma$ describes a cross-sectional area with a surface normal parallel to $\mathbf{k}_\circ$, through which the number of particles dN that get scattered into the angle Ω flow per unit time. The total cross section

$$\sigma_{\rm tot} = \int_0^{4\pi} \left(\frac{d\sigma}{d\Omega}\right) d\Omega \tag{14}$$

is the total effective geometrical cross-sectional area of the incident beam that is intercepted and the particles therein deflected by the target object.
From Eq. (13) the scattered current density is $j_{\rm sc} = \frac{1}{r^2} j_{\rm in}(\frac{d\sigma}{d\Omega})$. On the other hand $j_{\rm sc}$ can be calculated directly employing the expression of the probability current density given by

$$\mathbf{j}_{sc}(\mathbf{r}) = -i\frac{\hbar}{2m}\left[\psi^*_{\rm sc}(\mathbf{r})\nabla\psi_{\rm sc}(\mathbf{r}) - \psi_{\rm sc}(\mathbf{r})\nabla\psi^*_{\rm sc}(\mathbf{r})\right] \simeq j_{\rm in}\frac{1}{r^2}|f(\Omega)|^2\,\hat{\mathbf{r}} + \mathcal{O}\left(\frac{1}{r^3}\right)\ldots \tag{15}$$

where as $\psi_{\rm sc} = \frac{1}{r}e^{ikr}f(\Omega)$ is the asymptotic scattering wave, Eq. (8). We explicitly inserted here the current density $\mathbf{j}_{\rm in} = \frac{\hbar\mathbf{k}}{m}$ to the incoming plane wave $\psi_{\rm in}$. Equating the two expressions gives the relation

$$I(\Omega) \propto \left(\frac{d\sigma}{d\Omega}\right) = |f(\Omega)|^2 \tag{16}$$

for the differential cross section. This expression relates the experimental quantity, the differential cross section, to the scattering amplitude, which characterizes the wave function at large distances from the target. It is the fundamental relation between scattering theory and scattering experiments.

3 Lippmann Schwinger Equation

Having established the basic concepts for the scattering problem, we turn now to the illustration of the physical ideas that underlie the scattering analysis using integral equation methods. We recall that we are looking for the solution of the stationary Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_{\mathbf{k}}(\mathbf{r}) = E\psi_{\mathbf{k}}(\mathbf{r}) \qquad \text{with} \qquad V(\mathbf{r}) = 0 \qquad \text{except } \mathbf{r} \in \text{ target region } \mathbb{T}, \tag{17}$$

that is consistent with the boundary condition (8) of an incident plane wave $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}$ and an emanating scattered wave. The energy E is determined by the energy of the incident plane wave $E_{\mathbf{k}} = \frac{\hbar^2}{2m}k^2$. By introducing the Green function $G_\circ$,

$$\left[\frac{\hbar^2}{2m}\nabla^2 + E\right] G_\circ(\mathbf{r}, \mathbf{r}'|E) = \delta(\mathbf{r} - \mathbf{r}'), \tag{18}$$

for the potential-free Schrödinger equation, the Schrödinger equation for $\psi_{\mathbf{k}}(\mathbf{r})$,

$$\left[\frac{\hbar^2}{2m}\nabla^2 + E\right] \psi_{\mathbf{k}}(\mathbf{r}) = V(\mathbf{r})\, \psi_{\mathbf{k}}(\mathbf{r}), \tag{19}$$

can be transformed into an integral equation

$$\psi'_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}(\mathbf{r}) + \int_{\mathbb{T}} d^3r' G_\circ(\mathbf{r}, \mathbf{r}'|E)\, V(\mathbf{r}')\, \psi'_{\mathbf{k}}(\mathbf{r}'), \tag{20}$$

in which the formal expression $V(\mathbf{r})\, \psi'_{\mathbf{k}}(\mathbf{r})$ is conceived as inhomogeneity of the differential equation (18). This integral equation is called the *Lippmann-Schwinger equation*. Hereby, $\psi_{\mathbf{k}}(\mathbf{r})$ is the above cited plane-wave solution of the potential-free Schrödinger equation. The index $\mathbf{k}$ in ψ' expresses the fact that this state has evolved from one that in the remote past was a plane wave of the particular wavevector $\mathbf{k}$. Obviously, in the limit of zero potential, $V(\mathbf{r}) \to 0$, the scattered and the incident wave are identical, $\psi'_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}(\mathbf{r})$.
The Green function $G_\circ(\mathbf{r}, \mathbf{r}'|E)$ is not uniquely determined by the Schrödinger equation (18). Also here the unique solution requires a boundary condition, which is chosen such, that the solution $\psi'_{\mathbf{k}}(\mathbf{r})$ describes outgoing scattered waves. The Green function $G_\circ(\mathbf{r}, \mathbf{r}'|E)$,

$$G_\circ(\mathbf{r}, \mathbf{r}'|E) = -\frac{2m}{\hbar^2}\frac{1}{4\pi}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \qquad \text{with} \qquad k = \sqrt{\frac{2m}{\hbar^2}\, E}, \tag{21}$$

describes then the stationary radiation of a particle of energy E, that is generated at $\mathbf{r}'$, by a spherical wave outgoing from the target. In other words, the Green function $G_\circ(\mathbf{r}, \mathbf{r}'|E)$ gives the amplitude of this wave at location $\mathbf{r}$ due to its generation by the source at $\mathbf{r}'$, under the condition that the wave is not further scattered during its propagation from $\mathbf{r}'$ to $\mathbf{r}$. By the Lippmann-Schwinger equation, the incident wave $\psi_{\mathbf{k}}(\mathbf{r})$ is superimposed with spherical waves emitted from scattering at position $\mathbf{r}'$ in the target. The amplitude of these scattered waves is proportional to the interaction potential $V(\mathbf{r}')$ and the amplitude of the total wave field $\psi'(\mathbf{r}')$ at that point.
Recalling our experimental set-up that the distance between target and detector is significantly larger than the size of the sample, for large distances between $\mathbf{r}$ and the scattering center $\mathbf{r}'$

it is useful to expand the Green function $G_\circ$ in powers of $\frac{r'}{r} \ll 1$ assuming that the extent of $\mathbf{r}'$ is restricted to the space of a small target or scattering volume, respectively, $\mathbf{r}' \in \mathbb{T}$. Approximating for $r' \ll r$

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{r} + \mathcal{O}\left(\frac{1}{r^2}\right) \qquad \text{and} \qquad |\mathbf{r}-\mathbf{r}'| \approx r - \hat{\mathbf{r}}\cdot\mathbf{r}' \qquad \text{with} \qquad \hat{\mathbf{r}} = \frac{\mathbf{r}}{r} \tag{22}$$

and inserting this into the relation (21) one obtains the asymptotic form, or far-field limit, respectively, of the Green function $G_\circ$,

$$G_\circ(\mathbf{r},\mathbf{r}'|E) = -\frac{2m}{\hbar^2}\frac{1}{4\pi}\frac{e^{ikr}}{r}e^{-ik\hat{\mathbf{r}}\cdot\mathbf{r}'} + \mathcal{O}\left(\frac{1}{r^2}\right). \tag{23}$$

Inserting this expression into the Lippmann-Schwinger equation (20) one obtains the asymptotic solution of the wave function $\psi'_{\mathbf{k}}(\mathbf{r})$ for large distances $\mathbf{r}$

$$\psi'_{\mathbf{k}}(\mathbf{r}) \simeq e^{i\mathbf{k}\mathbf{r}} + \frac{1}{r}e^{ikr} f_{\mathbf{k}}(\hat{\mathbf{r}}), \tag{24}$$

which is exactly the boundary condition (8) we conjectured from Huygens' principle, whereas the scattering amplitude $f(\hat{\mathbf{r}}) = f(\theta,\phi)$ is given by the integral,

$$f_{\mathbf{k}}(\hat{\mathbf{r}}) = -\frac{2m}{\hbar^2}\frac{1}{4\pi}\int d^3r'\, e^{-i\mathbf{k}'\mathbf{r}'}V(\mathbf{r}')\,\psi'_{\mathbf{k}}(\mathbf{r}') = -4\pi\frac{\hbar^2}{2m}\,T(\mathbf{k}',\mathbf{k}) \tag{25}$$

that can be interpreted as a transition-matrix element from the scattering state described by $\psi'_{\mathbf{k}}(\mathbf{r}')$ to the scattered state at far distances, which is a plane-wave state described by $\mathbf{k}' = k\cdot\hat{\mathbf{r}}$, the wave vector of the scattered wave in the direction of the detector, which is known in the experiment. $T(\mathbf{k}',\mathbf{k})$ is referred to as the T matrix or transition amplitude, a quantity proportional to the scattering amplitude. Due to the far-field approximation (22) the scattering pattern $f_{\mathbf{k}}(\hat{\mathbf{r}})$ is independent of the distance between target and detector, depending only on the angles to the detector from the target. In optics this is known as the Fraunhofer diffraction and in this context approximation (23) is also referred to as the Fraunhofer approximation of the Green function.

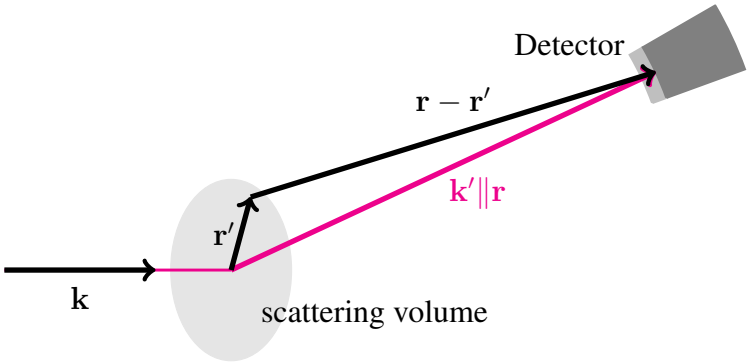


Fig. 3: *Scattering geometry for the calculation of the far-field limit at the detector. In the Fraunhofer approximation, we assume that* $|\mathbf{r}| \gg |\mathbf{r}'|$.

4 Born Approximation

Note that in the Lippmann-Schwinger equation (20) the wave function $\psi'(\mathbf{k})$ appears both on the left and right hand side. In a general case, there is no simple way to find exact solutions of the Lippmann-Schwinger equation. The form of the Lippmann-Schwinger equation provides a natural but approximate means that can be used for any potential, under the proper conditions, to proceed by an iterative procedure. At zeroth order in V, the scattering wave function is specified by the unperturbed incident plane wave,

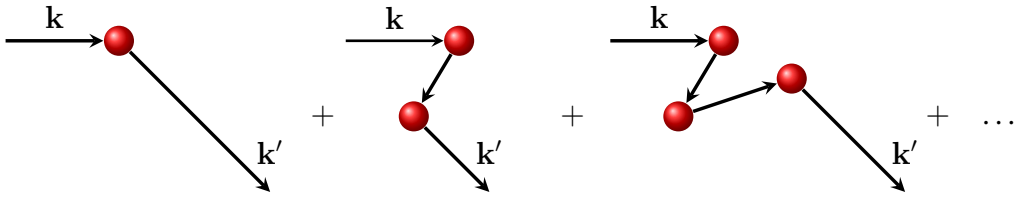
$$\psi'^{(0)}_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}. \tag{26}$$

Then one can iterate the Lippmann-Schwinger equation (20) according to the rule

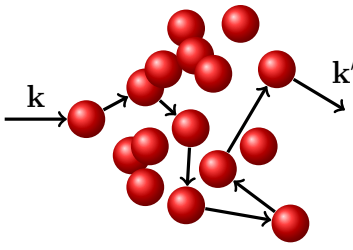
$$\psi'^{(n+1)}_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} + \int d^3r' G_\circ(\mathbf{r},\mathbf{r}'|E)\, V(\mathbf{r}')\, \psi'^{(n)}_{\mathbf{k}}(\mathbf{r}') \tag{27}$$

that results in the *Born expansion* of the wave function in powers of the interaction potential V written here in a symbolic form[4]

$$\begin{aligned}
\psi'_{\mathbf{k}} &= \psi'^{(0)}_{\mathbf{k}} + \psi^{(1)}_{\mathbf{k}} + \psi^{(2)}_{\mathbf{k}} + \psi^{(3)}_{\mathbf{k}} + \cdots + && (28)\\
&= \psi'^{(0)}_{\mathbf{k}} + G_\circ V \psi'^{(0)}_{\mathbf{k}} + G_\circ V G_\circ V \psi'^{(0)}_{\mathbf{k}} + G_\circ V G_\circ V G_\circ V \psi'^{(0)}_{\mathbf{k}} + \cdots + && (29)\\
&= (1 + G_\circ T)\, \psi'^{(0)}_{\mathbf{k}} \quad \text{with} \quad T = V + G_\circ V + \cdots + = \frac{1}{1 - VG_\circ} && (30)
\end{aligned}$$



A term-by-term convergence of this series is in general not guaranteed and depends on the potential and the energy of the incident particle, even though the final expression is always valid. Physically, an incoming particle undergoes a sequence of *multiple scattering* events from the potential. The first term in the series expansion (29) describes single scattering processes of the incident wave, while the following terms describe then scattering processes of successively higher order. Rarely are higher-order terms calculated analytically, since the complications then become so great that one might as well use a numerical method to obtain the exact solution if this is possible at all. Thus, only the first iteration of the series is taken into account, i.e. only single scattering, and the T matrix is approximated by the potential matrix $V(\mathbf{k}',\mathbf{k})$,



$$T(\mathbf{k}',\mathbf{k}) \simeq V(\mathbf{k}',\mathbf{k}). \tag{31}$$

This first order term, in which the exact wave function $\psi'_{\mathbf{k}}(\mathbf{r}')$ in the integral kernel is replaced by the plane wave $e^{i\mathbf{kr}'}$ is the *first Born approximation* and typically abbreviated as the *Born approximation*.[5] This approximation is most useful when calculating the scattering amplitude.

[4] Please note that $\psi^{(n)}_{\mathbf{k}} = (G_\circ V)^n \psi'^{(0)}_{\mathbf{k}}$. This is different from definition $\psi'^{(n)}_{\mathbf{k}}(\mathbf{r})$ in (27).

[5] It should not be confused with the Born-Oppenheimer approximation.

In first Born approximation the general equation for the scattering amplitude (25) reads then

$$f_{\mathbf{k}}^{(1)}(\hat{\mathbf{r}}) = -\frac{2m}{\hbar^2}\frac{1}{4\pi}\int d^3r'\, e^{-i\mathbf{k}'\mathbf{r}'}V(\mathbf{r}')\, e^{i\mathbf{k}\mathbf{r}'} = -\frac{2m}{\hbar^2}\frac{1}{4\pi}V(\mathbf{Q}) \quad \text{with} \quad \mathbf{Q}=\mathbf{k}-\mathbf{k}', \tag{32}$$

with $V(\mathbf{Q})$ denoting the Fourier transform of the potential with the momentum transfer $\mathbf{Q}$.[6] $V(\mathbf{Q})$ can be interpreted as a transition-matrix describing the transition from the incoming plane-wave of state $\mathbf{k}$ into the outgoing plane-wave state $\mathbf{k}'$ due to the action of the potential expressed in the reciprocal space at scattering angle $\mathbf{Q}$. From (16) follows then the differential cross section

$$\left(\frac{d\sigma}{d\Omega}\right)^{(1)} = \left(\frac{2m}{\hbar^2}\right)^2 \frac{\pi}{2}|V(\mathbf{Q})|^2. \tag{33}$$

The physics behind the 1st Born approximation is provided by the assumption that the incoming wave scatters only *once* inside the target potential before forming the scattered wave $\psi^{(1)}$. This is the concept behind the *kinematic theory of scattering*, that simplifies the interpretation of the scattering experiment substantially. For example, for the case of elastic scattering that we assumed all the time during the derivations, energy is conserved $|\mathbf{k}|^2 = |\mathbf{k}'|^2$, all possible scattering vectors are placed on the so-called *Ewald sphere* with radius $|\mathbf{k}|$. The length of the *scattering vector* $\mathbf{Q}$ is then given by

$$Q(\Omega) = |\mathbf{Q}| = 2\,|\mathbf{k}|\,\sin\frac{1}{2}\Omega = \frac{4\pi}{\lambda}\,\sin\frac{1}{2}\Omega \qquad \text{with} \qquad \Omega=(\theta,\phi)\sphericalangle(\mathbf{k}',\mathbf{k}). \tag{34}$$

Note, and this is the essence of the Ewald-sphere, that this shows that the differential cross section (33) does not depend on scattering angle and beam energy independently, but on a single parameter through the combination $Q = 2k\,\sin\frac{1}{2}\Omega$. By using a range of energies, $\mathbf{k}$, for the incoming particles, this dependence can be used to test whether experimental data can be well described by the Born approximation. A very common use of the Born approximation is, of course, in reverse. Having found $\frac{d\sigma}{d\Omega}$, experimentally, a reverse Fourier transform can be used to obtain the form of the potential.

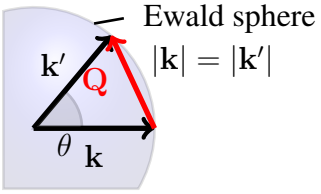


Fig. 4: *The Ewald sphere.*

4.1 Example of Born Approximation: Central Potential

For a centrally symmetry potential, $V(\mathbf{r}) = V(r)$, we can make some progress with the matrix element integral (32) if we choose a polar coordinate system with $\mathbf{Q}$ along the z-axis, so that $\mathbf{Q}\cdot\mathbf{r} = Qr\cos\theta$. Then, the scattering amplitude in Born approximation $f^{(1)}(\theta)$ is written after some manipulations in the form

$$f_k^{(1)}(\theta) = -\frac{2m}{\hbar^2}\frac{1}{Q}\int_0^\infty V(r)\, r\, \sin Qr\, dr \tag{35}$$

and is seen to be independent of ϕ due to the cylindrical symmetry of the problem at hand and all scattering vectors are placed on an Ewald-circle. An example, is the *Rutherford* scattering or Coulomb scattering, respectively, where a charged particle with charge $Z_1\mathrm{e}$ impinges on an

[6] It would be mathematically more correct to denote the Fourier transformation $\mathcal{F}$ of $V(\mathbf{r})$ by a different function name e.g. $\widetilde{V}(\mathbf{Q}) = \mathcal{F}[V(r)]$. To avoid incomprehension of reading due to unduly complicated notation we replace $\widetilde{V}(\mathbf{Q})$ by $V(\mathbf{Q})$.

other charged particle with charge $Z_2\mathrm{e}$ under the action of a Coulomb potential, which results into the scattering amplitude

$$f_k^{(1)}(\theta) = -\frac{2m}{\hbar^2}\frac{Z_1 Z_2}{Q^2}\mathrm{e}^2 = -\frac{1}{4}\frac{Z_1 Z_2}{\sin^2\frac{1}{2}\theta}\mathrm{e}^2\frac{1}{E} \tag{36}$$

and the differential cross section,

$$\left(\frac{d\sigma}{d\theta}\right)^{(1)} = \left|\frac{1}{4}\frac{Z_1 Z_2}{\sin^2\frac{1}{2}\theta}\mathrm{e}^2\right|^2 \frac{1}{E^2} \tag{37}$$

known as the *Rutherford formula*. Due to the long-range nature of the Coulomb scattering potential, the boundary condition on the scattering wave function does not apply. We can, however, address the problem by working with the screened (Yukawa) potential, $V(r) = \frac{Z_1 Z_2}{r}e^{-\kappa r}$, leading to $f^{(1)} \propto \frac{1}{Q^2+\kappa^2}$ and taking $\kappa \to 0$, which leads then to the Rutherford formula (37). Accidentally, the first Born approximation gives the correct result of the differential cross section for the Coulomb potential.

4.2 Example of Born Approximation: Square Well Potential

Consider scattering of particles interacting via a spherical three dimensional (3D) square well potential $V(r) = V_\circ$ for $r \le R_\circ$ and zero outside ($V(r) = 0$ for $r > R_\circ$). The integral (35) for the scattering amplitude required here is then

$$f_k^{(1)}(\theta) = \frac{2m}{\hbar^2}\frac{1}{Q}\int_0^{R_\circ} V_\circ\, r\, \sin Qr\, dr = \frac{2m}{\hbar^2}\frac{1}{Q}V_\circ\left[\frac{\sin Qr - Qr\cos Qr}{Q^2}\right]_0^{R_\circ} \tag{38}$$

[7] and whence to the differential cross section

$$\left(\frac{d\sigma}{d\theta}\right)^{(1)} = \left(\frac{2m}{\hbar^2}\frac{V_\circ}{Q}\right)^2 R_\circ^2\, j_1^2(QR_\circ) \simeq \left(\frac{2m}{\hbar^2}\frac{V_\circ}{Q}\right)^2 \begin{cases} \frac{1}{9}\left(1-\frac{1}{5}Q^2R_\circ^2\right) & \text{for low } E, \quad kR_\circ < 1 \\ \frac{R_\circ^2}{Q^2} & \text{for high } E, \quad kR_\circ > 1 \end{cases} . \tag{39}$$

From integrating over θ and ϕ the low and high energy limits for the total cross section are

$$\sigma(E\to\infty) = \pi\left(\frac{2m}{\hbar^2}\right)^2\left(\frac{V_\circ R_\circ^3}{kR_\circ}\right)^2 \qquad \sigma(E\to 0) = \sigma(E\to\infty)\frac{8}{9}\left(k^2R_\circ^2 - \frac{2}{5}k^4R_\circ^4 + \cdots\right). \tag{40}$$

The two examples illustrate some general features of scattering in the Born approximation:
(i) Born approximation is based on perturbation theory, so it works best for high energy particles.
(ii) At high energy, the scattering amplitude and the cross section are inversely proportional to the energy ($E = \hbar^2k^2/2m$). E.g. both become smaller and the scattering weaker with increasing energy. This is a general phenomenon, if no bound states appear in the vicinity of the energy. This can be seen best by inspecting the Fourier transformed Green function $G_\circ(k|E) \propto 1/(E - \frac{h^2k^2}{2m})$ that is inverse proportional to the energy.
(iii) Scattering depends on square of the interaction potential, e.g. $V_\circ^2$, so both attractive and

[7] $j_0(Qr) = \sin Qr/Qr$ is the spherical Bessel function for angular momentum $\ell = 0$. Radial integration leads to Bessel function $j_1(Qr)$.

repulsive potentials behave the same.
(iv) The dependence on the energy of the incident beam k and scattering angle θ arises only through the combination $Q = 2k \sin \frac{\theta}{2}$. Thus as energy increases, the scattering angle θ is reduced and the scattered beam becomes more peaked in the forward direction.
(v) Angular dependence depends on the range of the potential $R_\circ$ but not on the strength $V_\circ$.
(vi) The total cross section depends on both range $R_\circ$ and depth $V_\circ$ of the potential.

4.3 Validity of first Born Approximation

This raises the practical questions (i) under which conditions the Born expansion converges and (ii) whether the first term is a good approximation. In the Born approximation the T matrix is approximated by the potential matrix V. This will not work if the denominator $|1 - VG_\circ|$ in (30) is small or zero. This is the situation at low energy, when the energy of the incoming beam coincides with bound states of the potential. Then, the Born approximation is invalid and the Born expansion will not converge. The solution to this problem is provided by the *dynamical scattering theory* discussed in Chapter A3. According to (30) the Born approximation $T \simeq V$ is equivalent to the condition

$$\left| \int\int_{\mathbb{T}} d^3r \, d^3r' \, V(\mathbf{r}) G_\circ(\mathbf{r}, \mathbf{r}'|E) V(\mathbf{r}') \right| \ll 1 \,. \tag{41}$$

At the same time this condition determines the radius of convergence of the Born series with respect to the strength of the potential. This condition means that the first Born approximation is valid and the Born series converges if the potential is sufficiently weak and the approximation improves as the energy is increased. Concerning the question whether the first term is itself a good approximation to the wave function, a convenient, although nonrigorous, criterion can be obtained by requiring that the first-order correction to the wave function be small compared to the incident wave in the region of the potential, i.e. $|\psi'^{(1)}_{\mathbf{k}}(\mathbf{r})| \ll |\psi'^{(0)}_{\mathbf{k}}(\mathbf{r})|$ which results to

$$\frac{2m}{\hbar^2} \frac{1}{4\pi} \left| \int_{\mathbb{T}} d^3r' \, e^{-ikr'} V(\mathbf{r}') \, e^{-i\mathbf{k}\mathbf{r}'} \right| \ll 1. \tag{42}$$

For the above introduced spherical 3D square well potential $V(r \leq R_\circ) = V_\circ$ and $V(r > R_\circ) = 0$, this implies

$$\left| \frac{mV_\circ}{\hbar^2 k^2} \left(e^{ikR_\circ} \sin kR_\circ - kR_\circ \right) \right| \ll 1. \tag{43}$$

or

$$\frac{m}{\hbar^2} |V_\circ| R_\circ^2 \ll 1 \qquad \text{for low energies } kR_\circ < 1 \tag{44}$$

$$\frac{m}{\hbar^2} |V_\circ| R_\circ \frac{1}{k} \ll 1 \qquad \text{for high energies } kR_\circ > 1. \tag{45}$$

Since a bound state for this potential exists when $\frac{m}{\hbar^2}|V_\circ|R_\circ^2 \gtrsim 1$, as said above, the Born approximation will not be valid at low energies if the potential is so strong that it has a bound state. On the other hand criterion (45) can be satisfied for any potential by going to sufficiently high energy. When we square criterion (45) and multiply it by the geometrical cross section $\sigma_{\text{geo}} = \pi R_\circ^2$, criterion (45) reads

$$\pi \left(\frac{2m}{\hbar^2} \right)^2 \left(\frac{V_\circ R_\circ^3}{kR_\circ} \right)^2 \ll \pi R_\circ^2 \qquad \Longleftrightarrow \qquad \sigma_{\text{tot}} \ll \sigma_{\text{geo}}. \tag{46}$$

and provides a hand-waving criterion when the potential is sufficiently weak so that the Born approximation gives reliable results: If the ratio of the scattering cross section and the geometrical extension of the potential is small, $u := \frac{\sigma_{\rm tot}}{\sigma_{\rm geo}} \ll 1$, the Born approximation can be used. For x-ray and neutron scattering, the scattering cross sections amount to a few 10^{-24} cm^2, the cross-sectional area per atom is of the order of several 10^{-16} cm^2. This results indeed in a very small potential strength of $u \sim 10^{-8} \div 10^{-7}$ for scattering on different atoms: that means, the Born approximation is justified and the easy-to-interpret kinematic interpretation of scattering results is sufficient. The arguments become invalid for the nuclear scattering of neutrons by individual nuclei as the cross-sectional area of a nucleus is eight orders of magnitude smaller and the scattering cross section and the geometrical gross section can be of similar size and the potential strength u can be even larger than 1, $u > 1$. Due to the strong Coulomb interaction potential, the probability for multiple scattering processes of electrons in solids is extremely high, making the interpretation of electron diffraction experiments very difficult. Although in neutron and x-ray scattering, the first Born approximation is almost always adequate, even for neutrons and x-rays, the kinematic scattering theory can break down, for example in the case of Bragg scattering from large nearly perfect single crystals. In this case as in the case of electron scattering the wave equation has to be solved exactly under the boundary conditions given by the crystal geometry. This is then called the dynamic scattering theory discussed in Chapter A3. For simple geometries, analytical solutions can be obtained. Other examples where the Born series do not converge are neutron optical phenomena like internal total reflection in a neutron guide, or grazing-incidence small-angle neutron scattering (GISANS). The same holds for x-ray scattering for example in combination with grazing-incidence small-angle x-ray scattering (GISAXS) experiments. The *grazing-incidence small-angle scattering (GISAS)* techniques and their application will be discussed in Chapter D2. The theoretical analysis makes use of the *distorted-wave Born approximation (DWBA)*.

4.4 Distorted-Wave Born Approximation (DWBA)

In the previous Section we discussed that the Born approximation is accurate if the scattered field is small, compared to the incident field, in the scatterer. The scatterer is treated as a perturbation to free space or to a homogeneous medium, and the incident wave is a plane wave. When this smallness criteria is not met, it is often possible to generalize the idea of the Born approximation, which is frequently referred to as the distorted-wave Born approximation (DWBA). In generalization to the Born approximation, the free space zero-potential, $V_\circ(\mathbf{r}) = 0$, is replaced by a non-trivial reference potential $V_1(\mathbf{r})$ to which the scattered field $\psi'^1_{\mathbf{k}}$ is known analytically, numerically, e.g. due to the solution of the Lippmann-Schwinger equation (20),

$$\psi'^1_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} + \int_{\mathbb{T}} d^3r' G_\circ(\mathbf{r},\mathbf{r}'|E)\, V_1(\mathbf{r}')\, \psi'^1_{\mathbf{k}}(\mathbf{r}'), \tag{47}$$

or experimentally. The interaction of interest V

$$V(\mathbf{r}) = V_1(\mathbf{r}) + \delta V(\mathbf{r}) \qquad \text{with} \qquad |\delta V| \ll |V_1| \tag{48}$$

is treated as a perturbation δV to the reference system V_1. In the distorted-wave Born approximation, the scattering field $\psi'_{\mathbf{k}}(\mathbf{r})$ due to the potential V is then determined applying the Born approximation

$$\psi'_{\mathbf{k}}(\mathbf{r}) = \psi'^1_{\mathbf{k}}(\mathbf{r}) + \int d^3r' G_1(\mathbf{r},\mathbf{r}'|E)\, \delta V(\mathbf{r}')\, \psi'^1_{\mathbf{k}}(\mathbf{r}') \tag{49}$$

to the description of the scattering of the incident wave $\psi'^1_{\mathbf{k}}(\mathbf{r})$, the so-called "distorted" wave, due the perturbative potential $\delta V(\mathbf{r})$. The "distorted" incident wave, is the outgoing-wave solution of

$$\left[\frac{\hbar^2}{2m}\nabla^2 - V_1(\mathbf{r}) + E\right]\psi'^1_{\mathbf{k}}(\mathbf{r}) = 0, \tag{50}$$

that is supposed to be known, and $G_1(\mathbf{r},\mathbf{r}'|E)$ is the corresponding Green function with the outgoing boundary condition for the same potential,

$$\left[\frac{\hbar^2}{2m}\nabla^2 - V_1(\mathbf{r}) + E\right]G_1(\mathbf{r},\mathbf{r}'|E) = \delta(\mathbf{r}-\mathbf{r}'). \tag{51}$$

In analogy to the potential-free case (19), the difference to the reference system that appears in the Schrödinger equation, $\delta V(\mathbf{r})\psi'^1_{\mathbf{k}}(\mathbf{r})$, can be considered as inhomogeneity that constitutes a Lippmann-Schwinger equation with $\psi'^1_{\mathbf{k}}(\mathbf{r})$ as homogeneous solution. The Born approximation to this equation is given by Eq. (49).
To satisfy the boundary conditions we must also require that the "distorted" wave function behaves in the asymptotic limit as plane wave plus an outgoing wave

$$\psi'^1_{\mathbf{k}}(\mathbf{r}) \stackrel{r\to\infty}{\longrightarrow} e^{i\mathbf{kr}} + \frac{1}{r}e^{ikr}f^1_{\mathbf{k}}(\theta,\phi), \tag{52}$$

where, as in (25)

$$f^1_{\mathbf{k}}(\theta,\phi) = -\frac{2m}{\hbar^2}\frac{1}{4\pi}\int d^3r'\, e^{-i\mathbf{k}'\mathbf{r}'}V_1(\mathbf{r}')\,\psi'^1_{\mathbf{k}}(\mathbf{r}'). \tag{53}$$

This is simply the scattering amplitude for the potential $V_1(\mathbf{r})$, as if it were the only potential present, assumed to be known. The total scattering amplitude $f_{\mathbf{k}}(\theta,\phi)$ is

$$f_{\mathbf{k}}(\theta,\phi) = f^1_{\mathbf{k}}(\theta,\phi) + \delta f_{\mathbf{k}}(\theta,\phi) \tag{54}$$

where $\delta f_{\mathbf{k}}(\theta,\phi)$ is calculated in the Born approximation ($\psi'_{\mathbf{k}}(\mathbf{r}) \simeq \psi'^1_{\mathbf{k}}(\mathbf{r})$)

$$\delta f_{\mathbf{k}}(\theta,\phi) \simeq -\frac{2m}{\hbar^2}\frac{1}{4\pi}\int d^3r'\,\psi'^{1(-)*}_{\mathbf{k}'}(\mathbf{r}')\delta V(\mathbf{r}')\,\psi'^1_{\mathbf{k}}(\mathbf{r}'). \tag{55}$$

The scattering amplitude describes the scattering strength of an outgoing spherical wave. By inspection of Eq. (53) one finds that the first wave function of the integrand is a plane wave $e^{-i\mathbf{k}'\mathbf{r}'}$, whose negative sign in the exponent represents an incoming plane wave. According of the standard definition of plane waves we can write $e^{-i\mathbf{k}'\mathbf{r}'} = \psi^{(-)*}_{\mathbf{k}}$, where $(-)$ denotes the incoming boundary condition. Quite in the same way $\psi'^{1(-)*}_{\mathbf{k}'}(\mathbf{r}')$ is the known incoming wave function corresponding to the reference potential V_1.
Clearly Eq. (55) will be a good approximation if $\delta V(\mathbf{r})$ is sufficiently small, so that the additional scattering that is generated does not significantly modify the wave function. Some example in which this method is useful include scattering in which $\delta V(\mathbf{r})$ may be the spin-orbit interaction or a perturbation due to many-particle excitations, atomic scattering where $\delta V(\mathbf{r})$ may be a deviation from the Coulomb potential or from a Hartree average potential, or in case of scattering at a magnetic superlattice where $V_1(\mathbf{r})$ contains the scattering at the nuclei or electron charge distribution plus the interaction to an average magnetization, and $\delta V(\mathbf{r})$ describes the interaction to the modulated magnetic structure of the superlattice. The DWBA is at place analyzing grazing-incidence small-angle scattering (GISAS) experiments to resolve the magnetic structure of superlattices [6].

5 Method of Partial Wave Expansion

The differential equation formulation of scattering provides additional insights that are not readily apparent from the integral equation discussed in the previous section. Many potentials in nature are spherically symmetric, or nearly so, and thus for simplicity, here we will focus on the properties of a centrally symmetric potential, $V(r)$, where the scattering wave function, $\psi'(\mathbf{r})$ (and indeed that scattering amplitudes, $f(\theta)$) must be symmetrical about the axis of incidence, and hence independent of the azimuthal angle, ϕ. The method of partial wave expansion is inspired by the observation that a plane wave $\psi_{\mathbf{k}} = e^{i\mathbf{kr}}$ can actually be written as a sum over spherical waves

$$\psi_{\mathbf{k}} = e^{i\mathbf{kr}} \;=\; e^{ikr\cos\theta} \;=\; \sum_{\ell=0}^{\infty}(2\ell+1)\, i^{\ell} j_{\ell}(kr) P_{\ell}(\cos\theta)\,, \tag{56}$$

known as the Rayleigh expansion. As we shall discuss in more detail below, the real function is a standing wave, made up of incoming and outgoing waves of equal amplitude. The radial functions $j_\ell(kr)$ appearing in the above expansion of a plane wave in its spherical components are the *spherical Bessel functions*, discussed below.
Generalizing this concept, if we define the direction of the incident wave $\mathbf{k}$ to lie along the z-axis, and θ denotes the scattering angle to the detector, $\theta = \sphericalangle(\mathbf{k}, \mathbf{r})$, then the azimuthal rotational symmetry of plane waves and the spherical potential around the direction of the ingoing wave ensures that the wave function can be expanded in a series

$$\psi(\mathbf{r}) = \psi(r,\theta) = \sum_{\ell=0}^{\infty} (2\ell+1) i^{\ell}\, R_{\ell}(r)\, P_{\ell}(\cos\theta) \tag{57}$$

of Legendre polynomials $P_\ell(\cos\theta) = \sqrt{\frac{4\pi}{2\ell+1}} Y_{\ell 0}(\theta)$, where $Y_{\ell m}$ denotes the spherical harmonics. Each term in the series is known as a *partial wave*, and is a simultaneous eigenfunction of the angular momentum operators $\mathbf{L}^2$ and L_z having eigenvalue $\hbar^2\ell(\ell+1)$, and 0, respectively. Following standard spectroscopic notation, $\ell = 0, 1, 2, \cdots$ are referred to as $s, p, d, \cdots$ waves. The *partial wave amplitudes*, f_ℓ are determined by the radial functions, $R_\ell(r)$, defined by

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} - v(r) + k^2\right] R_{\ell}(r,E) = 0 \qquad \text{with} \qquad V(r) = 0 \quad \text{except } r \in \mathbb{RT}, \tag{58}$$

where $v(r) = \frac{2m}{\hbar^2} V(r)$ represents the effective potential and k^2 refers to the energy of the incoming beam $k^2 = \frac{2m}{\hbar^2} E_{\mathbf{k}}$. The energy $E_{\mathbf{k}}$ can be chosen positive and equal to the kinetic energy of the projectile when it is far from the scattering center. The potential $V(r)$ or $v(r)$, respectively, will be assumed to vanish sufficiently rapidly with increasing r that it may be neglected beyond some finite radius, that defines the radial target region $\mathbb{RT}$ or scattering volume, respectively. We are looking for the solution of the stationary Schrödinger equation that is consistent with the boundary condition (8) of an incident plane wave $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}$ and an emanating spherical scattered wave. Beyond the range of the potential, i.e. r outside the radial target region $\mathbb{RT}$, the $R_\ell(r,E)$ may be expressed in terms of the solutions of the potential free radial differential equation

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} + k^2\right] R_{\ell}(r,E) = 0. \tag{59}$$

This is a differential equation of 2nd order which has two linearly independent solutions at each energy E, known as the spherical Bessel function

$$R_\ell(r,E) = j_\ell(kr) \quad \text{with} \quad j_\ell(kr) \stackrel{rk\ll 1}{\longrightarrow} \frac{z^\ell}{(2l+1)!!} \quad \text{and} \quad j_\ell(kr) \stackrel{rk\gg 1}{\longrightarrow} \frac{1}{kr}\sin\left(kr-\ell\frac{\pi}{2}\right) \tag{60}$$

and the spherical Neumann function

$$R_\ell(r,E) = n_\ell(kr) \quad \text{with} \quad n_\ell(kr) \stackrel{rk\ll 1}{\longrightarrow} \frac{(2\ell-1)!!}{z^{\ell+1}} \quad \text{and} \quad n_\ell(kr) \stackrel{rk\gg 1}{\longrightarrow} \frac{1}{kr}\cos\left(kr-\ell\frac{\pi}{2}\right), \tag{61}$$

whereas j_ℓ and n_ℓ show a regular and irregular solutions, respectively, in the origin $r = 0$ and $n!! = n(n-2)(n-4)\cdots 1$. That means any solution $R_\ell(r)$ of the radial Schrödinger equation (59) can be expressed at a given energy E for r outside $\mathbb{RT}$ as linear combination of j_ℓ and n_ℓ or in the form spherical Hankel functions

$$h_\ell^{(\pm)}(kr) = n_\ell(kr) \pm i\, j_\ell(kr) \stackrel{rk\gg 1}{\longrightarrow} \frac{1}{kr} e^{\pm i\left(kr-\ell\frac{\pi}{2}\right)}, \tag{62}$$

a different set of independent solutions that correspond to incident $(-)$ and emanating $(+)$ radial waves at large distances r. This holds also for the wave function of the incident beam before scattering expressed in terms of a plane wave

$$\psi_{\mathbf{k}} = e^{i\mathbf{kr}} = \sum_{\ell=0}^{\infty}(2\ell+1)\, i^\ell j_\ell(kr) P_\ell(\cos\theta) = \frac{i}{2}\sum_{\ell=0}^{\infty}(2\ell+1)\, i^\ell \left(h_\ell^{(-)}(kr) - h_\ell^{(+)}(kr)\right) P_\ell(\cos\theta), \tag{63}$$

that can be recast according to Rayleigh into incoming and outgoing spherical Hankel functions. After scattering, the incoming spherical wave $h_\ell^{(-)}$ is unaffected by the scattering process, while the outgoing wave $h_\ell^{(+)}$ is modified by a herewith introduced quantity,

$$S_\ell(k) \quad \text{or} \quad S_\ell(E) = e^{i2\delta_\ell(E)}, \tag{64}$$

the *partial wave scattering matrix*, subject to the constraint $|S_\ell(k)| = 1$ following from the conservation of particle flux (current density times area). $\delta_\ell(E)$ is the *phase shift* (the name becomes clear below as the phase difference between incoming and outgoing wave). For scattering processes where the net flux of particles is zero, the phase shift is real, and thus only the phase and not the amplitude of the outgoing spherical wave is affected but the presence of the potential. The wave after scattering $\psi'(\mathbf{r})$ reads then

$$\psi'_{\mathbf{k}}(\mathbf{r}) = \psi'_{\mathbf{k}}(r,\cos\theta) = \frac{i}{2}\sum_{\ell=0}^{\infty}(2\ell+1) i^\ell \left(h_\ell^{(-)}(kr) - S_\ell(k)\, h_\ell^{(+)}(kr)\right) P_\ell(\cos\theta) \tag{65}$$

$$= \sum_{\ell=0}^{\infty}(2\ell+1) i^\ell \left(j_\ell(kr) + T_\ell(k)\, h_\ell^{(+)}(kr)\right) P_\ell(\cos\theta) \quad r \notin \mathbb{RT}. \tag{66}$$

The first term in the parenthesis proportional to j_ℓ sums up according to the Rayleigh expansion (56) to the incoming plane wave, the second describes the outgoing spherical wave multiplied by a *partial wave scattering amplitude* $f_\ell(k)$ or the *partial wave transition matrix element* $T_\ell(k)$

$$T_\ell(k) = \frac{1}{2i}(S_\ell(k)-1) = e^{i\delta_\ell(k)}\sin\delta_\ell(k) = \frac{1}{\cot\delta_\ell - i} = k f_\ell(k) \tag{67}$$

due to the presence of the interaction potential. The wave function after scattering takes the asymptotic form

$$\psi'_{\mathbf{k}}(\mathbf{r}) \simeq e^{i\mathbf{k}r} + \sum_{\ell=0}^{\infty}(2\ell+1)i^{\ell}\,T_{\ell}(k)\,\frac{1}{kr}e^{i\left(kr-\ell\frac{\pi}{2}\right)}P_{\ell}(\cos\theta) \;= e^{i\mathbf{k}r} + \frac{1}{r}e^{ikr}f_{\mathbf{k}}(\theta). \tag{68}$$

consistent with the scattering boundary condition (8) where the scattering amplitude $f_{\mathbf{k}}(\theta)$ can be related to the partial wave scattering amplitude and the phase shift as

$$f_{\mathbf{k}}(\theta) = \sum_{\ell=0}^{\infty}(2\ell+1)f_{\ell}(k)P_{\ell}(\cos\theta)\,. \tag{69}$$

Making use of the identity $\int d\Omega P_{\ell}(\cos\theta)P_{\ell'}(\cos\theta) = \frac{4\pi}{2\ell+1}\delta_{\ell\ell'}$ and the definition of the total cross section (14) one obtains

$$\sigma_{\rm tot}(k) = \int |f_{\mathbf{k}}(\theta)|^2 d\Omega = \sum_{\ell=0}^{\infty}\sigma_{\ell}(k) = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)|T_{\ell}(k)|^2 = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\sin^2\delta_{\ell}(k)\,. \tag{70}$$

The total cross section is additive in the contribution of the $\sigma_{\ell}(k)$ of each partial wave. This does not mean, though, that the differential cross-section for scattering into a given solid angle is a sum over separate ℓ values, no the different components interfere. It is only when all angles are integrated over, that the orthogonality of the Legendre polynomials guarantees that the cross-terms vanish.

Notice that the scattering cross-section for particles in angular momentum state ℓ is upper bounded by

$$\sigma_{\ell}(k) \le \frac{4\pi}{k^2}(2\ell+1)\,, \tag{71}$$

which is four times the classical cross section for that partial wave impinging on, e.g. a hard sphere: Imagine semi-classically particles in an annular area, with the angular momentum $L = rp$, but $L = \hbar\ell$ and $p = \hbar k$ so $\ell = rk$. Therefore, the annular area corresponding to angular momentum between ℓ and $\ell+1$ has inner and outer radii, ℓ/k and $(\ell+1)/k$, respectively, and therefore the area $\frac{\pi}{k^2}(2\ell+1)$. The quantum result is essentially a diffractive effect.

The maximal contribution is obtained for the phase shifts $\delta_{\ell}(k) = (n+\frac{1}{2})\pi$, with $n = 0, \pm1, \pm2, \cdots$. For these energies $E_{\mathbf{k}}$, resonant scattering occurs if in addition $\delta_{\ell}(k)$ changes rapidly. On the other hand, for energies leading to phase shifts $\delta_{\ell}(k) = n\pi$ with $n = 0, \pm1, \pm2, \cdots$, the scattering amplitude and the cross section vanish.

Since for the imaginary part of the partial wave scattering amplitude (67) holds

$$\Im f_{\ell}(k) = \frac{1}{k}\sin^2\delta_{\ell}(k) = k|f_{\ell}(k)|^2 \quad \text{or more simply} \quad \Im\frac{1}{f_{\ell}(k)} = -k \tag{72}$$

and the Legendre polynomial at unity are always unity, $P_{\ell}(1) = 1$ for $\forall\ell$, and apply this to equation (69) we find that

$$\Im f_{\mathbf{k}}(0) = \frac{k}{4\pi}\,\sigma_{\rm tot}(k)\,, \tag{73}$$

a relation known as the *optical theorem*. It is a direct consequence of the flux conversation in elastic scattering and says for example that the scattering amplitudes are complex valued quantities.

Comparing equation (57) with equation (66) and replacing the definition of the partial wave transition matrix T_ℓ by the phase shift given in (67) we can write the radial wave function $R_\ell(r,E)$ after scattering outside the target region, $r \notin \mathbb{RT}$, as

$$\begin{aligned} R_\ell(r,E) &= j_\ell(kr) + h_\ell^{(+)}(kr)\, e^{i\delta_\ell(k)} \sin\delta_\ell(k) \qquad \text{for} \quad r \notin \mathbb{RT} \quad (74)\\ &= e^{i\delta_\ell(k)} \left(\cos\delta_\ell(k) j_\ell(kr) + \sin\delta_\ell(k) n_\ell(kr)\right) \quad (75)\\ &\simeq e^{i\delta_\ell(k)} \frac{1}{kr} \left(\cos\delta_\ell(k) \sin\left(kr - \ell\frac{\pi}{2}\right) + \sin\delta_\ell(k) \cos\left(kr - \ell\frac{\pi}{2}\right)\right) \quad (76)\\ &\simeq e^{i\delta_\ell(k)} \frac{1}{kr} \sin\left(kr - \ell\frac{\pi}{2} + \delta_\ell(k)\right) \qquad \text{for} \quad kr \gg 1\,. \quad (77) \end{aligned}$$

In the asymptotic limit, the radial incoming wave $j_\ell(kr) \simeq \frac{1}{kr}\sin\left(kr - \ell\frac{\pi}{2}\right)$ and the scattered wave differ by just a phase $\delta_\ell(k)$ known as the scattering phase, which gives $\delta_\ell(k)$ the name phase shift, as well as a phase phase factor $e^{i\delta_\ell(k)}$.
I would like to end this section with remarking that the scattering and transition matrices S, T, respectively, describe the scattering at different boundary condition. The scattering matrix describes the scattering from the incoming spherical wave into an outgoing spherical wave, while the transition matrix describes scattering from an incoming plane wave into an emanating spherical wave. The scattering matrix contains all the scattered and the unscattered states and the matrix elements are unity without scattering. The T matrix contains only the scattered states and it has only zero valued matrix elements in the absence of scattering.

5.1 The Born Approximation for Partial Waves

From the boundary condition (8) and the solution of the Lippmann-Schwinger equation (see Section 3) in far-field limit

$$\psi'_{\mathbf{k}}(\mathbf{r}) \simeq e^{i\mathbf{kr}} - \frac{2m}{\hbar^2}\frac{1}{4\pi}\int d^3r'\, e^{-i\mathbf{k}'\mathbf{r}'} V(\mathbf{r}')\, \psi'_{\mathbf{k}}(\mathbf{r}') \quad (78)$$

we obtained the respective definition of scattering amplitude $f(\theta)$ (25). On inserting expression (56) and (57) for the plane wave and the wave function after scattering, respectively, and integrating over the angle $d\Omega'$ one yields the radial Lippmann-Schwinger equation for the far field limit,

$$R_\ell(r,k) \simeq j_\ell(kr) - \int_{\mathbb{RT}} r'^2 dr'\, j_\ell(kr')\, v(r')\, R_\ell(r',k) \quad (79)$$

and an explicit formulation of the partial scattering amplitudes

$$f_\ell(\theta) = \frac{1}{k} e^{i\delta_\ell(k)} \sin\delta_\ell(k) = -\int_{\mathbb{RT}} r'^2 dr'\, j_\ell(kr')\, v(r')\, R_\ell(r',k) \quad (80)$$

which provides an elegant procedure to calculate the phase shift. We recall that in the first Born Approximation the exact wave function $R_\ell(r',k)$ in the integral kernel is replaced by the plane wave represented by the Bessel function $j_\ell(kr')$ and the partial-wave Born approximation of the scattering matrix and the transition matrix, respectively, reads

$$f_\ell^{(1)}(\theta) = \frac{1}{k} T_\ell^{(1)}(k) = \frac{1}{k} e^{i\delta_\ell^{(1)}(k)} \sin\delta_\ell^{(1)}(k) = -\int_{\mathbb{RT}} r'^2 dr'\, j_\ell(kr')\, v(r')\, j_\ell(kr') \approx \frac{1}{k}\delta_\ell^{(1)}(k) \quad (81)$$

with an approximate expression for the phase shift in Born approximation $\delta_\ell^{(1)}(k)$ valid for small phase shifts (the only place where the Born approximation is valid).
Concerning the Distorted-Wave Born Approxmiation (DWBA) one can also perform a partial-wave analysis of Eq. (54) to obtain an approximate expression for the phase shift. This result is

$$e^{i\delta_\ell^{(1)}(k)} \sin\delta_\ell^{(1)}(k) = e^{i\delta_\ell^1(k)} \sin\delta_\ell^1(k) - \int_{\mathbb{RT}} r'^2 dr' \, \left(R_\ell^1(kr')\right)^2 \, v(r') \,, \tag{82}$$

where $\delta_\ell^1(k)$ are the phase shifts to the nontrivial reference potential V_1, and the Bessel function representing the plane waves are replaced by the exact radial scattering solution $R_\ell^1(kr)$.

5.2 Low Energy Scattering: Scattering Phases and Scattering Length

From (81) follows that the sign of $\delta_\ell(k)$ is determined by the sign of the potential. For an attractive potential, the phase shift $\delta_\ell > 0$ is positive and the phase shift is negative, $\delta_\ell < 0$, for a repulsive potential. At large distances r, the zeros of $R_\ell(r,k) \simeq \frac{1}{kr}\sin\left(kr - \ell\frac{\pi}{2} + \delta_\ell(k)\right)$ are at $r_0 = \frac{1}{k}\left(n\pi + \ell\frac{\pi}{2} - \delta_\ell(k)\right)$. Positive (negative) $\delta_\ell(k)$ relate to an inward (outward) shift of the nodes. That means, for an attractive potential, the probability of a particle to stay in the potential range becomes greater, so that the wave function is drawn into the potential range: the nodes shift inwards, i.e. $\delta_\ell > 0$. For a repulsive potential the wave function is squeezed out of the potential range. In consequence the nodes move to the outside, i.e. $\delta_\ell > 0$.
If $kr' \ll 1$ or $\lambda \gg r'$, respectively, we shall be able to approximate the Bessel function $j_\ell(kr') \simeq \frac{1}{(2\ell+1)!!}(kr')^\ell$ and one obtains the simple estimate

$$\delta_\ell(k) \approx \left(\frac{1}{(2\ell+1)!!}\right)^2 k^{2\ell+1} \int_{\mathbb{RT}} dr' r'^{2\ell+2} \, v(r) \tag{83}$$

for the scattering phase. For low energies and high angular momenta the scattering phases $\delta_\ell(k)$ behave proportional to $\delta_\ell(k) \propto k^{2\ell+1}$. In particular one expects that only *s-wave scattering* ($\ell = 0$) survives for $k \to 0$ since the cross section scales as

$$\sigma_\ell(k) = \frac{4\pi}{k^2}(2\ell+1)\sin^2\delta_\ell(k) \propto k^{4\ell} \,. \tag{84}$$

When a slow particle scatters off a short ranged scatterer it cannot resolve the structure of the object since its de Broglie wavelength λ is very long, larger than the scatterer. The idea is that then it should not be important what precise potential $V(\mathbf{r})$ one scatters off, but only how the potential looks at long length scales. At very low energy the incoming particle does not see any structure, therefore to lowest order one has only a spherical symmetric outgoing wave, the so called s-wave scattering (angular momentum $\ell = 0$). At higher energies one also needs to consider p and d-wave (l = 1,2) scattering and so on.
Although exact at all energies, the partial wave method is most useful for dealing with scattering of low energy particles. This is because for slow moving particles to have large angular momentum ($\hbar kb$) they must have large impact radii b. Classically, particles with impact radius larger than the range of the potential miss the potential. Thus, for scattering of slow-moving particles we need only to consider a few partial waves, all the others are unaffected by the potential ($\delta_\ell \approx 0$). Thus at a given incoming momentum, $\hbar k$, we can determine how many terms in the partial wave expansion to consider from $\hbar k b_{\max} \approx \ell_{\max}\hbar$, where $b_{\max}$ is the maximum impact parameter for classical collision, i.e. the range of the potential $\mathbb{RT}$. Since the angular

variation of the Legendre polynomial for the angular momentum $\ell = 0$ is $P_{\ell=0}(\cos\theta) = 1$, the s-wave scattering is isotropic, consistent with the thought of averaging over the potential. Since in practical applications the expansion into Legendre polynomials has to be truncated for higher ℓ values, since otherwise the effort becomes too large, the partial wave analysis is primarily a method of approximation for low energies. It generally requires an exact (numerical) solution of the radial equations, since the Born approximation fails at low energies in general. Thus partial waves and the Born approximation are complementary methods, good for slow and fast particles, respectively.
At energy $E \to 0$, the radial Schrödinger equation for $rR_{\ell=0}$ away from the potential becomes $\frac{d^2}{dr^2} rR_{\ell=0} = 0$ with a straight line solution $rR_{\ell=0} = (r - a_s)$.[8] For the s-wave solution the approximation (77) becomes exact and the radial wave function $rR_0(r,k) = \sin(kr + \delta_0(k)) \stackrel{k\to 0}{\longrightarrow} k\left(r + \frac{1}{k}\delta_\ell(k)\right)$ can only become a straight line in r, if $\delta_0(k)$ is itself linear in k for sufficiently small k. Then $\delta_0(k) = -ka_s$, a_s being the point at which the extrapolated external wave function intersects the axis (maybe at negative r). So, as k goes to zero, the term

$$\lim_{k\to 0} k\cot\delta_0(k) = -\frac{1}{a_s} \tag{85}$$

dominates in the denominator of expression (67) where the scattering amplitude and the cross section take the expression

$$f_{\ell=0}(k\to 0) = -a_s \qquad \text{and} \qquad \sigma_{\ell=0}(k\to 0) = 4\pi a_s^2\,. \tag{86}$$

The parameter a_s of dimension length is called the *scattering length*. At low energies it determines solely the elastic cross section. This is a nontrivial construction from the potential itself and the wave function of the state. We see that for momenta much less than the inverse radius of the potential the scattering length is sufficient to describe all of the interactions. It is clear that by measuring the scattering length of a system alone we cannot reconstruct the potential uniquely. There are infinitely many different shapes, depths and ranges of potentials that will reproduce a single scattering length.

5.3 S-Wave Scattering at Square Well Potential

The properties of scattering phases are studied for the problem of quantum scattering from an attractive spherically symmetric three-dimensional (3D) square well potential $V(r) = -\frac{\hbar^2}{2m} v_0 \Theta(R_\circ - r)$. For convenience we write $v_0 = k_v^2$. The continuity condition of the wave function R_ℓ

$$R_\ell(r,k) = A_\ell(\kappa)\, j_\ell(\kappa r) \qquad \text{with} \qquad \kappa^2 = k^2 + k_v^2 \qquad \text{for} \quad r \le R_\circ \tag{87}$$
$$R_\ell(r,k) = e^{i\delta_\ell(k)} \left(\cos\delta_\ell(k)\, j_\ell(kr) + \sin\delta_\ell(k)\, n_\ell(kr)\right) \qquad \text{for} \quad r \ge R_\circ \tag{88}$$

and its derivative $R'_\ell = \frac{d}{dr} R_\ell$ at the boundary, $r = R_\circ$, where $A_\ell(\kappa)$ is a normalization constant. The wave function outside of $R_\circ$ is normalized to to incoming plane wave normalized to unity.

$$A_\ell(\kappa)\, j_\ell(\kappa R_\circ) = e^{i\delta_\ell(k)} \left(\cos\delta_\ell(k)\, j_\ell(kR_\circ) + \sin\delta_\ell(k)\, n_\ell(kR_\circ)\right) \tag{89}$$
$$\kappa\, A_\ell(\kappa)\, j'_\ell(\kappa R_\circ) = k\, e^{i\delta_\ell(k)} \left(\cos\delta_\ell(k)\, j'_\ell(kR_\circ) + \sin\delta_\ell(k)\, n'_\ell(kR_\circ)\right) \tag{90}$$

[8] Normalization constant A_0 is neglected for simplicity.

translates to the following relation for the phase shifts

$$\tan\delta_\ell(k) = -\frac{k\,j'_\ell(kR_\circ) - L_\ell(E)\,j_\ell(kR_\circ)}{k\,n'_\ell(kR_\circ) - L_\ell(E)\,n_\ell(kR_\circ)} = \frac{1}{\cot\delta_\ell(k)} \qquad \text{with} \qquad L_\ell(E) = \kappa\,\frac{j'_\ell(\kappa R_\circ)}{j_\ell(\kappa R_\circ)}, \tag{91}$$

where L_ℓ is the logarithmic derivative of the wave function at the potential boundary. Here $j'_\ell(x) = \frac{d}{dx} j_\ell(x)$ and similarly for $n'_\ell(x)$.

Hard Sphere Potential

The simplest case is the scattering at a hard sphere potential:

$$v_\circ(r) = \infty \quad \text{for} \quad r < R_\circ \qquad \text{and} \qquad v_\circ(r) = 0 \quad \text{for} \quad r \geq R_\circ\,. \tag{92}$$

Since then the wave function at the boundary $r = R_\circ$ vanishes, it follows that $L_\ell(v_\circ \to \infty) \to \infty$, and the phase shift reduces to

$$\tan\delta_\ell(k) = -\frac{j_\ell(kR_\circ)}{n_\ell(kR_\circ)} \qquad s\text{-wave}: \quad \tan\delta_0(k) = -\frac{(\sin kR_\circ)/kR_\circ}{(\cos kR_\circ)/kR_\circ} = -\tan kR_\circ \tag{93}$$

so that $\delta_0(k) = kR_\circ$. Thus, the s-wave radial wave function for $r > R_\circ$ takes the form

$$R'_0(r) = e^{-ikR_\circ}\left(\cos kR_\circ \frac{\sin kr}{kr} - \sin kR_\circ \frac{\cos kr}{kr}\right) = e^{-ikR_\circ}\frac{1}{kr}\sin k(r - R_\circ). \tag{94}$$

The corresponding radial wave-function for the incident wave takes the form $R_0(r) = \frac{1}{kr}\sin kr$. It is clear that the actual $\ell = 0$ radial wave function is similar to the incident wave function, except that it is phase-shifted by $kR_\circ$. According to (86) the total s-wave cross-section of the hard wall potential yields then $\sigma^\infty_{\ell=0} = 4\pi R_\circ^2$, four times the geometric cross-section $\sigma_{\rm geo} = \pi R_\circ^2$ (i.e., the cross-section for classical particles bouncing off a hard sphere of radius $R_\circ$). However, low energy scattering implies relatively long wave-lengths, so we do not necessarily expect to obtain the classical result in this limit. Recall that the s-wave scattering is a good approximation to the low-energy scattering.

Consider the high energy limit $kR_\circ \gg 1$. At high energies, all partial waves up to $\ell_{\max} = kR_\circ$ contribute significantly to the scattering cross-section. With so many ℓ values contributing, it is legitimate to replace $\sin^2\delta_\ell$ in the expression (70) for the partial wave cross section by its average value $1/2$ and thus (for comparison we include also the low energy result, i.e. $kR_\circ \ll 1$)

$$\sigma^\infty_{\rm tot}(kR_\circ \ll 1) \simeq \sigma^\infty_{\ell=0} = 4\pi R_\circ^2 \qquad \text{and} \qquad \sigma^\infty_{\rm tot}(kR_\circ \gg 1) \simeq 2\pi R_\circ^2\,. \tag{95}$$

This is twice the classical result $\sigma_{\rm geo} = \pi R_\circ^2$, which is somewhat surprising, since we might expect to obtain the classical result in the short wave-length limit. For hard sphere scattering, incident waves with impact parameters less than $R_\circ$ must be deflected. However, in order to produce a "shadow" behind the sphere, there must be scattering in the forward direction (recall the optical theorem) to produce destructive interference with the incident plane-wave. In fact, the interference is not completely destructive, and the shadow has a bright spot in the forward direction. The effective cross-section associated with this bright spot is $\pi R_\circ^2$ which, when combined with the cross-section for classical reflection, $\pi R_\circ^2$, gives the actual cross-section of $2\sigma_{\rm geo}$.

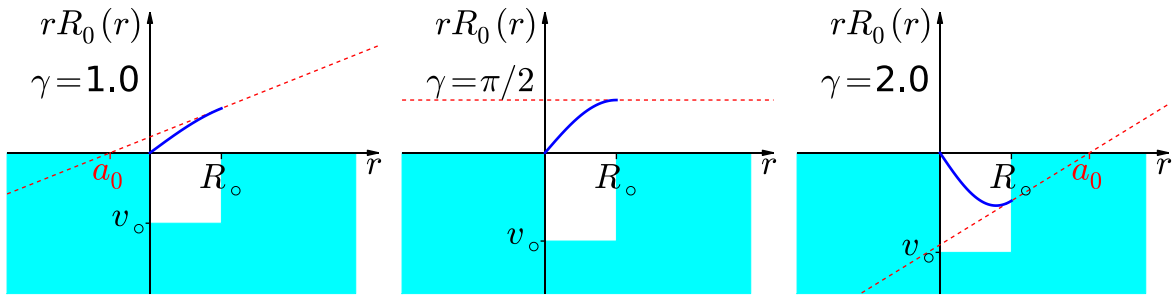


Fig. 5: *Radial scattering wave function, $rR_0(r)$, for three-dimensional square well potential of radius $R_\circ$ for $kR_\circ = 0.1$ and $\gamma = k_v R_\circ = 1$ (left), $\pi/2$ (middle) and 2 (right). Note that the scattering length, a_0 changes from negative to positive as system passes through bound state.*

Soft Sphere Potential

For a potential with finite scattering strength $V(r) = -\frac{\hbar^2}{2m} v_\circ(r)$ the logarithmic derivative of the scattering potential L_ℓ is finite. Analogously to (93) for s-wave scattering, Eq. (93) is simplified to

$$\tan\delta_0(k) = \frac{k\tan(\kappa R_\circ) - \kappa\tan(kR_\circ)}{\kappa + k\tan(kR_\circ)\tan(\kappa R_\circ)}. \tag{96}$$

Then, unless $\tan(\kappa R_\circ) = \infty$, an expansion at low energy (small k) yields $\delta_0(k) \simeq kR_\circ \times$ $\times \left(\frac{\tan(\kappa R_\circ)}{\kappa R_\circ} - 1\right)$, and the $\ell = 0$ partial cross section,

$$\sigma_{\ell=0}(k) = \frac{4\pi}{k^2}\sin^2\delta_0(k) = \frac{4\pi}{k^2}\frac{1}{1+\cot^2\delta_0(k)} \simeq \frac{4\pi}{k^2}\delta_0^2(k) = 4\pi R_\circ^2\left(\frac{\tan(\kappa R_\circ)}{\kappa R_\circ} - 1\right) \tag{97}$$

From this result we find that, when $\frac{\tan(\kappa R_\circ)}{\kappa R_\circ} = 1$, the scattering cross-section vanishes. An expansion in small k obtains,

$$k\cot\delta_0(k) = -\frac{1}{a_0} + \frac{1}{2}r_0 k^2 + \cdots, \tag{98}$$

where $a_0 = \left(1 - \frac{\tan(k_v R_\circ)}{k_v R_\circ}\right) R_\circ$, defines the scattering length a_0 or a_s, respectively, and r_0 is the effective range of the interaction that is obtained from the Taylor expansion of (96) for small $kR_\circ$. At low energies, $k \to 0$, the scattering cross section, $\sigma_0 = 4\pi a_0^2$ (see above) is fixed by the scattering length alone. If $|k_v R_\circ| \ll 1$, a_0 is negative. As $k_v R_\circ$ is increased, when $k_v R_\circ = \pi/2$, both a_0 and σ_0 diverge there is said to be a zero energy resonance. This condition corresponds to a potential well that is just able to support an s-wave bound state at zero energy. If $k_v R_\circ$ is further increased, a_0 turns positive as it would be for an effective repulsive interaction until $k_v R_\circ = \pi$ when $\sigma_0 = 0$ and the process is repeated with the appearance of a second bound state at $k_v R_\circ = 3/2$, and so on. Since the scattering state must be orthogonal to all bounded states of the potential $V(r)$, the radial wave function of the scattering state must be orthogonal to all radial wave functions of bound states at equal angular momentum ℓ.[9] Consider the situation of a potential that supports at a given energy two bound states of s character. Then the scattering

[9]The orthogonality to states of different ℓ is automatically taken care of by orthogonality conditions of the angular part of the wave function.

wave should also have two additional nodes, which moves the scattering phase δ_ℓ about 2π. This is quantitatively expressed by the *Levinson Theorem*

$$\delta_\ell(k=0) = N_\ell \pi \tag{99}$$

where N_ℓ is the number of bound states at given angular momentum ℓ. If $\delta_\ell(k)$ is increasing rapidly through an odd multiple of $\pi/2$, $\sin^2\delta_\ell = 1$ the ℓ-th partial cross-section takes its maximum value and the cross-section exhibits a narrow peak as a function of energy and there is said to be a resonance. The analysis leads to the Breit-Wigner formula that goes beyond the scope of this lecture.

5.4 Nuclear Scattering Length

Two fundamental interactions govern the scattering of neutrons by an atomic system and define the neutron scattering cross-section measured in an experiment. The residual strong interaction, also known as the nuclear force, gives rise to scattering by the atomic nuclei (nuclear scattering). The electromagnetic interaction of the neutron's magnetic moment with the sample's internal magnetic fields gives rise to magnetic scattering. The latter is neglected in this chapter. The nuclear force is not weak as it is responsible for holding together protons and neutrons in the nucleus. However, it has extremely short range, 10^{-13} cm to 10^{-12} cm, comparable with the size of the nuclei, and much smaller than the typical distances in solids and much smaller than the typical neutron's wavelength. Consequently, away from the conditions of the resonance neutron capture, the probability of a neutron being scattered by an individual nucleus is very small. To describe the neutron's interaction with the atomic system in which the typical distances are about 1 Å (10^{-8} cm), the nuclear scattering length operator $\mathbf{b}_\mathrm{N}$ can be effectively treated as a δ-function in the coordinate representation

$$\mathbf{b}_\mathrm{N} = b_\mathrm{N}\delta(\mathbf{r} - \mathbf{R}_\mathrm{N})\,, \tag{100}$$

where $\mathbf{r}$ is a coordinate of a neutron and $\mathbf{R}_\mathrm{N}$ is that of a nucleus. Alternatively, in the momentum representation it is just a number (for the nucleus fixed at the origin), $\mathbf{b}_\mathrm{N}(\mathbf{Q}) = b_\mathrm{N}$, independent of the incident neutron's wave vector and of the wave-vector transfer, $\mathbf{Q}$. This again indicates that the applicability of such treatment is limited to neutrons whose wavelength is large enough compared to the size of the nuclei. In the Born approximation, Eq. (100) for the scattering length would correspond to the neutron-nucleus interaction,

$$V(\mathbf{r}, \mathbf{R}_\mathrm{N}) = -4\pi\frac{\hbar^2}{2m_\mathrm{n}} b_\mathrm{N}\delta(\mathbf{r} - \mathbf{R}_\mathrm{N}) \tag{101}$$

generally known as the Fermi pseudopotential [7, 8]. In Eqs. (100) and (101), the scattering length refers to the fixed nucleus. Usually, it is treated as a phenomenological parameter that is determined experimentally [9].

6 Scattering from a Collection of Scatterers

Finally, after having considered so far only the scattering at a single site with the target potential $V(\mathbf{r})$ placed at position $\mathbf{R}_\tau = \mathbf{0}$, prior to closing this chapter we shall relate these derivations

to the scattering phenomena in solid state systems with the potential composed of an assembly of targets

$$V(\mathbf{r}) = \sum_\tau v_\tau(\mathbf{r} - \mathbf{R}_\tau) \tag{102}$$

centered at a collection of sites $\mathbf{R}_\tau$. Inserting this into the Lippmann-Schwinger equation (20), replacing the integration variable $\mathbf{r}'$ by a vector $\mathbf{r}'_\tau \in \mathbb{T}_\tau$ within the target τ and the center-of-gravity-vector $\mathbf{R}_\tau$, $\mathbf{r}'_\tau + \mathbf{R}_\tau$, and taking into consideration that the free-space Green function $G_\circ(\mathbf{r}, \mathbf{r}'|E)$ depends only on $\mathbf{r} - \mathbf{r}'$, the Lippmann-Schwinger equation for many potentials can be written as

$$\psi'_\mathbf{k}(\mathbf{r}) = \psi_\mathbf{k}(\mathbf{r}) + \sum_\tau \int_{\mathbb{T}_\tau} d^3r'_\tau G_\circ(\mathbf{r} - \mathbf{R}_\tau, \mathbf{r}'_\tau|E)\, V(\mathbf{r}'_\tau)\, \psi'_\mathbf{k}(\mathbf{r}'_\tau + \mathbf{R}_\tau). \tag{103}$$

Approximating the Green function by its far-field asymptotic form (23) and considering that in the far-field solution $\psi'_\mathbf{k}(\mathbf{r}'_\tau + \mathbf{R}_\tau) \simeq e^{i\mathbf{k}\mathbf{R}_\tau}\psi'_\mathbf{k}(\mathbf{r}'_\tau)$, behind which is the Huygens' principle where the wave function generated at different sites share a phase relation, which becomes exact in the limit of the Born approximation, one obtains the asymptotic solution of the wave function $\psi'_\mathbf{k}(\mathbf{r})$ for large distances $\mathbf{r}$

$$\psi'_\mathbf{k}(\mathbf{r}) \simeq e^{i\mathbf{k}\mathbf{r}} + \sum_\tau \frac{1}{|\mathbf{r} - \mathbf{R}_\tau|} e^{ik|\mathbf{r}-\mathbf{R}_\tau|} e^{i\mathbf{k}\mathbf{R}_\tau} f_\mathbf{k}(\widehat{\mathbf{r} - \mathbf{R}_\tau}) \tag{104}$$

Since in the far-field approximation (22), $r \gg R_\tau$, $k|\mathbf{r} - \mathbf{R}_\tau| \simeq k(r - \hat{\mathbf{r}} \cdot \mathbf{R}_\tau) = kr - \mathbf{k}' \cdot \mathbf{R}_\tau$, we shall rewrite the asymptotic solution of the wave function for scattering at many potentials as

$$\psi'_\mathbf{k}(\mathbf{r}) \simeq e^{i\mathbf{k}\mathbf{r}} + \frac{1}{r} e^{ikr} F(\mathbf{Q}) \qquad \text{with} \qquad F(\mathbf{Q}) = \sum_\tau P_\tau(\mathbf{Q})\, f_{\tau\mathbf{k}}(\hat{\mathbf{r}}_\tau) \qquad \text{and} \qquad \mathbf{Q} = \mathbf{k}' - \mathbf{k}\,. \tag{105}$$

where the *static structure factor* (or *structure factor* for short) $F(\mathbf{Q})$ describes the way in which an incident beam is scattered by the atoms of a solid state system, taking into account the different scattering power of the elements through the term $f_{\tau\mathbf{k}}(\hat{\mathbf{r}}_\tau)$ also called *atomic form factor*

$$f_{\tau\mathbf{k}}(\hat{\mathbf{r}}_\tau) = -\frac{2m}{\hbar^2}\frac{1}{4\pi} \int_{\mathbb{T}_\tau} d^3r'_\tau\, e^{-i\mathbf{k}'\mathbf{r}'_\tau} V(\mathbf{r}'_\tau)\, \psi'_\mathbf{k}(\mathbf{r}'_\tau) \qquad \text{with} \qquad \mathbf{r}_\tau = \mathbf{r} - \mathbf{R}_\tau \tag{106}$$

that depends only on the potential and not on the position of the atom. The atomic form factor, or scattering power, of an element depends on the type of radiation considered. Since the atoms are spatially distributed, there will be a difference in phase when considering the scattered amplitude from two atoms. This phase difference is taken into account by the phase factor

$$P_\tau(\mathbf{Q}) = e^{i\mathbf{Q}\cdot\mathbf{R}_\tau}\,, \tag{107}$$

which depends only on the position of the atoms and is completely independent of the scattering potential. So in total the structure factor separates the interference effects from the scattering within an target from the interference effects arising from scattering from different targets. Thus, the scattering intensity and the differential cross section are proportional to the square of the structure factor

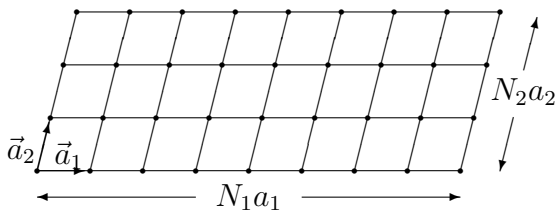
$$I(\mathbf{Q}) \propto \left(\frac{d\sigma}{d\Omega}\right) = |F(\mathbf{Q}|^2 = \left|\sum_\tau P_\tau(\mathbf{Q})\, f_{\tau\mathbf{k}}(\hat{\mathbf{r}}_\tau)\right|^2\,. \tag{108}$$

If we ignore spin degrees of freedom, so that we do not have to worry whether an electron does or does not flip its spin during the scattering process, then at low energies the scattering amplitude $f(\theta)$ of particles from a cluster of atoms or a crystal becomes independent of angle (s-wave) and maybe described by the scattering length b_τ for atom τ, i.e. $f_{\tau\mathbf{k}}(\hat{\mathbf{r}}_\tau) = b_\tau$. Then, the differential cross section simplifies to

$$\left(\frac{d\sigma}{d\Omega}\right) = \left|\sum_\tau P_\tau(\mathbf{Q})\, b_\tau\right|^2 . \tag{109}$$

If we consider scattering from a periodic crystal lattice, all atoms are same, i.e. have the same nuclear number (and we consider here also all nuclei as identical), thus $b_\tau = b$ for all atoms τ. Then, we are left with the differential cross section

$$\left(\frac{d\sigma}{d\Omega}\right) = N|b|^2 S(\mathbf{Q}) \qquad \text{with} \qquad S(\mathbf{Q}) = \frac{1}{N}\left|\sum_\tau e^{i\mathbf{Q}\mathbf{R}_\tau}\right|^2 . \tag{110}$$



$S(\mathbf{Q})$ is the form factor of the lattice, a quantity closely related to the pair density of a solid. For simplicity we consider scattering at a finite Bravais lattice. The lattice points are spanned by the lattice vector $\mathbf{R}_\tau = \mathbf{R}_\mathbf{m} = \boldsymbol{A}\mathbf{m}$, where $\boldsymbol{A}$ is the Bravais matrix consistent with the primitive vectors $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ of the three dimensional lattice, and $\mathbf{m} = (m_1, m_2, m_3) \in \mathbb{N}^3$ with $0 \le m_i \le (N_i - 1)$. The reciprocal lattice is defined by the matrix $\boldsymbol{B}$, that is orthogonal to $\boldsymbol{A}$, $\boldsymbol{B}^T\boldsymbol{A} = 2\pi\mathbf{1}$. The reciprocal lattice vectors are given as $\mathbf{G}_\mathbf{h} = \boldsymbol{B}\mathbf{h}$ with $\mathbf{h} = (h_1, h_2, h_3) \in \mathbb{Z}^3$. The transfer of scattering momentum vector is expressed by $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3) = \boldsymbol{B}\kappa$, with $\kappa = (\kappa_1, \kappa_2, \kappa_3) \in \mathbb{R}^3$. The scattering amplitude is calculated analytically summing up the geometrical series

$$\begin{aligned}
\sum_\mathbf{m} e^{i\mathbf{Q}\mathbf{R}_\mathbf{m}} &= \sum_\mathbf{m} e^{i2\pi\kappa\mathbf{m}} = \sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1}\sum_{m_3=0}^{N_3-1} e^{i2\pi\kappa_1 m_1} e^{i2\pi\kappa_2 m_2} e^{i2\pi\kappa_3 m_3} \\
&= \frac{1-e^{i2\pi\kappa_1 N_1}}{1-e^{i2\pi\kappa_1}} \cdot \frac{1-e^{i2\pi\kappa_2 N_2}}{1-e^{i2\pi\kappa_2}} \cdot \frac{1-e^{i2\pi\kappa_3 N_3}}{1-e^{i2\pi\kappa_3}} \\
&= e^{i2\pi(\kappa_1(N_1-1)/2+\kappa_2(N_2-1)/2+\kappa_3(N_3-1)/2)} \cdot \frac{\sin N_1\pi\kappa_1}{\sin\pi\kappa_1} \cdot \frac{\sin N_2\pi\kappa_2}{\sin\pi\kappa_2} \cdot \frac{\sin N_3\pi\kappa_3}{\sin\pi\kappa_3}
\end{aligned}$$

giving the scattered intensity

$$I(\mathbf{Q}) \propto \left(\frac{d\sigma}{d\Omega}\right) = |b|^2 \frac{\sin^2\frac{1}{2}N_1\mathbf{Q}_1\mathbf{a}_1}{\sin^2\frac{1}{2}\mathbf{Q}_1\mathbf{a}_1} \cdot \frac{\sin^2\frac{1}{2}N_2\mathbf{Q}_2\mathbf{a}_2}{\sin^2\frac{1}{2}\mathbf{Q}_2\mathbf{a}_2} \cdot \frac{\sin^2\frac{1}{2}N_3\mathbf{Q}_3\mathbf{a}_3}{\sin^2\frac{1}{2}\mathbf{Q}_3\mathbf{a}_3} . \tag{111}$$

taking into account that $2\pi\kappa_i = \mathbf{Q}_i\mathbf{a}_i$, for $i = 1, 2, 3$. The dependence of the scattering intensity on the scattering vector $\mathbf{Q}$ is given by the so-called *Laue function*, which separates according to the three Bravais vectors. One factor along one lattice direction a is plotted in Fig. 6.
The main peaks are the *Bragg reflections*. They occur at integer κ, $\kappa = (\kappa_1, \kappa_2, \kappa_3) \in \mathbb{Z}^3$, i.e. at reciprocal lattice vectors $\mathbf{Q} = \mathbf{G}_\mathbf{n}$. At points of Bragg reflection the coherent interference of scattering waves of all atoms add up constructively so that the maximum intensity scales

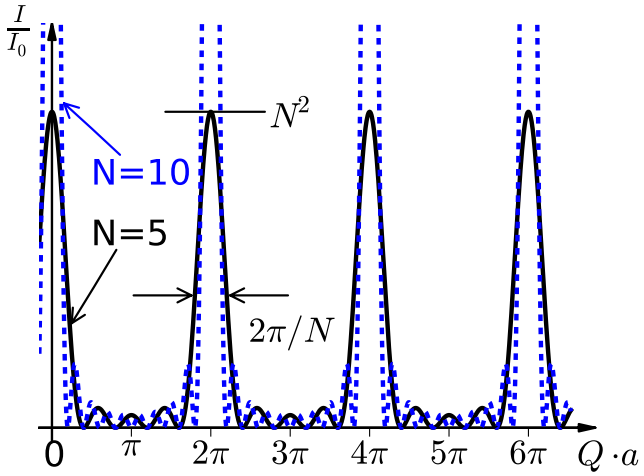


Fig. 6: *Laue function along the lattice direction a for a lattice with $N = 5$ and $N = 10$ periods.*

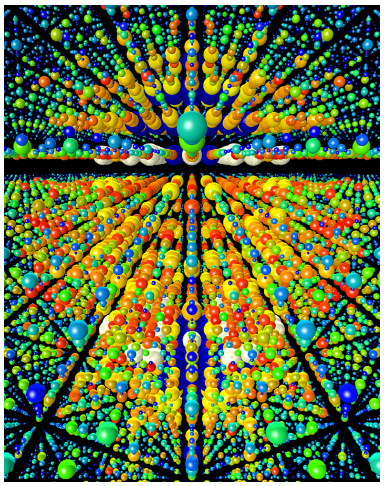


Fig. 7: *Three-dimensional rendering of x-ray diffraction data obtained from over 15 000 single nanocrystal diffraction snapshots of a protein complex [10].*

with the square of the number of periods N^2. This high intensity is the reason why the Born approximation can in general not be used to describe the scattering at Bragg peaks. At small deviations $\Delta\kappa$ from the exact Bragg condition the intensity drops fast,[10] so that the total intensity integrated over a small $\Delta\kappa$ region, $\Delta\kappa \simeq 1/N$, varies only $\simeq N$. The half width is given approximately by $\Delta Q = \frac{2\pi}{a}\frac{1}{N}$. The more periods contribute to coherent scattering, the sharper and higher are the main peaks. Between the main peaks, there are $N-2$ side maxima. With increasing number of periods N, their intensity becomes rapidly negligible compared to the intensity of the main peaks. From the position of these Bragg peaks in momentum space, the metric of the unit cell can be deduced (lattice constants a_1, a_2, a_3 in the three Bravais vector directions and unit cell angles α, β, γ). The width of the Bragg peaks is determined by the size of the coherently scattering volume $N = N_1N_2N_3$ and experimental factors (resolution) as well as details of the sample (size of crystallite, mosaic distribution, internal strains, etc.). For large N the form factor approaches a δ-function

$$\left(\frac{d\sigma}{d\Omega}\right) = N|b|^2 V_{\rm BZ}\sum_{\mathbf{h}}\delta(\mathbf{Q}-\mathbf{G_h})\,, \tag{112}$$

where $V_{\rm BZ}$ is the Brillouin-zone volume.

I shall conclude this chapter by mentioning that in 1914 the Nobel Prize in Physics was awarded to Max von Laue "for his discovery of the diffraction of x-rays by crystals" and in 1915 the Nobel Prize in Physics was awarded to Sir William Henry Bragg and William Lawrence Bragg "for their services in the analysis of crystal structure by means of x-rays". Since the pioneering

[10] For $\kappa \ll 1$ it follows that $\frac{\sin^2 N\pi\kappa}{\sin^2\pi\kappa} \approx \frac{\sin^2 N\pi\kappa}{\pi^2\kappa^2} \Longrightarrow \int_{-\infty}^{\infty} d\kappa\,\frac{\sin^2 N\pi\kappa}{\sin^2\pi\kappa} \approx \frac{N}{\pi}\int_{-\infty}^{\infty} dx\,\frac{\sin^2 x}{x^2} = N \Longrightarrow \bar{\kappa} \simeq 1/N$.

days of von Laue, Ewald, Knipping, Friedrich, the Bragg's, Compton, etc., diffraction experiments went a long way deciphering today the atomic arrangement of noncrystalline solids such as viruses as shown in a recent experiment [10] carried out at the Linac Coherent Light Source (LCLS), at SLAC National Accelerator Laboratory in California, USA) as shown in the diffraction image Fig. 7. Physical principles established 100 years ago and subsequent theoretical and experimental methods developed, reinvented and constantly brought to perfection contribute to day and will contribute in the future to the welfare of mankind.

Acknowledgements

I thank Dr. Phivos Mavropoulos and Prof. Peter Dederichs for discussions and Benedikt Schweflinghaus for his assistance in preparing the figures.

References

[1] J.J. Sakurai, *Modern Quantum Theory*, (Addison Wesley, 1984).

[2] Wu and Ohmura, *Quantum Theory of Scattering*, (Prentice Hall, 1962).

[3] A. S. Davidov, *Quantum Mechanics*, (Pergamon Press, 1991)

[4] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, (Brooks Cole, 1976)

[5] Chr. Huygens, Trait de la Lumiere (completed in 1678, published in Leyden in 1690)

[6] Sabrina Disch, Erik Wetterskog, Raphaël P. Hermann, German Salazar-Alvarez, Peter Busch, Thomas Brückel, Lennart Bergström, and Saeed Kamali, Nano Lett. **11**, 1651 (2011).

[7] G. L. Squires, *Introduction to the theory of thermal neutron scattering*, (Cambridge: Cambridge University Press, 1978); (New York: Dover Publications, 1996).

[8] E. Fermi, Ric. Sci. **7**, 13 (1936).

[9] E. Fermi, and L. Marshall, Phys. Rev. **71**, 666 (1947); Phys. Rev. **72**, 408 (1947).

[10] Henry N. Chapman *et al.*, Nature **470**, 73 (2011).

A 3 Scattering Theory: Dynamical Theory [1]

Yuriy Mokrousov
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

In order to describe the scattering of electrons and X-rays by crystals two theories are widely used. The first and the most common one is the so-called *kinematical theory*. Within the kinematical theory, it is assumed that the incoming wave of e.g. electrons or X-rays is scattered from the objects, which perform the scattering, e.g. atoms, only once. After such "once-scattering" the intensities of the scattered waves are added taking into account the phase differences of the scattered waves in order to form the intensities of the transmitted and reflected beams. What is neglected in the kinematical theory is the interaction of the "once-scattered" waves with each other and the matter. In terms of rigorous scattering theory, kinematical theory corresponds to the (1^{st}) *Born approximation*, discussed in detail in the proceeding's lecture by Prof. Stefan Blügel [1]. While kinematical theory works nicely in many cases, it is important to remember its limitations, namely, it works best when the crystals on which the incoming wave is scattered are small, thus, the interaction of the "once-scattered" wave with matter and other waves is negligible.

Very often, however, the diffraction of waves and particles on large crystals of very good crystalline quality is investigated, and the intensities of the transmitted and reflected beams are measured after the waves forming them traverse large regions of space filled with atoms. In this case the interaction between the scattered waves and atoms cannot be ignored anymore, and the so-called *multiple scattering* of waves inside the crystal, which technically corresponds to going beyond the Born approximation, has to be considered in order to explain observed phenomena which cannot be understood within the kinematic theory. Here, the situation can be essentially simplified if the perfect periodicity of the crystal is assumed. The foundations of the *dynamical scattering theory*, which goes beyond the kinematical scattering theory of diffraction along these lines, were set already by Darwin in 1914 [2]. On a more fundamental level the problem was treated by Ewald in 1917 [3] and later by Laue [4]. For electrons, the problem has been tackled further by Bethe [5]. Since then, many good books and reviews on the subject have been published, see e.g. the books of Zachariasen [6], Cowley [7] or Authier [8], or reviews by Slater [9], and Batterman and Cole [10]. In our manuscript, we mainly follow the review by Slater [9] and a beautiful review by P. H. Dederichs [11].

The present manuscript presents the simplest possible introduction into dynamical scattering, which is, however, mainly self-contained and can be understood without an advanced knowledge of quantum mechanics and solid state physics. The variety of effects which can be explained within the dynamical scattering theory is vast, and there are considerable differences for different types of radiation. There is, nevertheless, a common body of theory which lets us appreciate the main types of phenomena to be observed, and we present this in the simplest and most transparent form.

Below, summarized in Table I, we give an estimate of the most important for dynamical diffraction quantities. First, we give a typical energy E of neutrons, X-rays, electrons as well as for the case of low-energy-electron diffraction. Correspondingly, we also provide the characteristic wevelengths of different types of radiation λ. The most important for dynamical diffraction is the so-called *extinction length*, which is essentially the thickness of the crystal for which the kinematic theory breaks down. Note that for different types of radiation this thickness, denoted by d_{ext} in Table I, can change over five orders of magnitude. In particular, this means, that the dynamical theory is absolutely necessary for LEED, while it can be completely neglected for neutrons and it is not that important for X-rays, which can be more or less nicely described within the kinematic theory. Next, we provide the typical absorption length $1/\mu_0$, see sec-

	neutrons	X-rays	electrons	LEED
energy E	10 meV	10 keV	100 keV	100 eV
wavelength λ	1 Å	1 Å	0.05 Å	1 Å
extinction length d_{ext}	10^5 Å	10^4 Å	$10^2 - 10^3$ Å	5 Å
absorption length $1/\mu_0$	10^8 Å	10^5 Å	$10^3 - 10^4$ Å	10 Å
$1/\Delta\mu$	$> 10^8$ Å	$30 \cdot 10^5$ Å	$3 \cdot (10^3 - 10^4)$ Å	10 Å

Table 1: *Basic parameters for the diffraction of neutrons, electrons, X-rays and in the case low-energy electron diffraction (LEED).*

tion 4.1.1. We note, that the absorption length is essentially larger than the extinction length for neutrons and X-rays, while among all types of radiation neutrons are basically not absorbed at all by the medium. Finally, we give values for the absorption length in the case when the Bragg reflection is excited ($1/\Delta\mu$, where $\Delta\mu = \mu_0 \pm \mu_{\mathtt{G}}$, depending on the relative sign of μ_0 and $\mu_{\mathtt{G}}$, see section 4.1.1). This situation corresponds to the case of the so-called *anomalous transmission*, discussed in detail in section 4.1.1. We can observe from the Table I, that for the case of Bragg diffraction the absorption of the X-rays is reduced by a very large factor, known as the *Borrmann effect*, whereas the absorption of electrons is only slightly reduced.

2 Scattering of an electronic wave by a periodic perturbation: a preamble

We start with a simple picture of a wave of electrons, which is to be scattered, and which we can write as $e^{i(\omega_0 t - \mathbf{k}_0\mathbf{r})}$, where $\mathbf{k}_0$ is the propagation vector and $\mathbf{r} = (x, y, z)$ is the coordinate in real space. The wave which will do the scattering in a crystal we represent in a very simple way as $e^{-i\mathbf{G}\mathbf{r}}$, where $\mathbf{G}$ is some reciprocal vector of our crystal lattice. For simplicity, we assume in this section that the periodicity of the crystal is obeyed only along one direction, which is perpendicular to the surface of the crystal, onto which the wave $e^{i(\omega_0 t - \mathbf{k}_0\mathbf{r})}$ is incoming. Thus, $\mathbf{G} = (0, 0, G)$. This is a typical setup of the *Bragg diffraction* experiment, see Fig. 1. Generally speaking, the solution to this scattering experiment, that is, the decomposition of the total wavefunction in space into incident, scattered and the wavefunction inside the crystal can be obtained by finding a solution of the Schrödinger equation in $\mathbb{R}^3$, which satisfies certain boundary conditions. This solution must correspond to a constant energy $E = \hbar\omega_0$, meaning, that the wave inside the crystal, as well as incident and scattered waves have the same frequency. Let us first find the solution to the Schrödinger equation inside the crystal:

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V_1\cos(\mathbf{G}\mathbf{r})\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{1}$$

Owing to the Bloch's theorem, valid inside periodic crystals, every solution of this equation can be characterized by a certain wave vector $\mathbf{k}$ from the Brillouin zone of the crystal (reciprocal space) [12], and the solution itself can be represented as:

$$\psi(\mathbf{r}, t) = \sum_{n=-\infty}^{+\infty} A_n e^{-i[\omega_0 t - (\mathbf{k}+n\mathbf{G})\mathbf{r}]}. \tag{2}$$

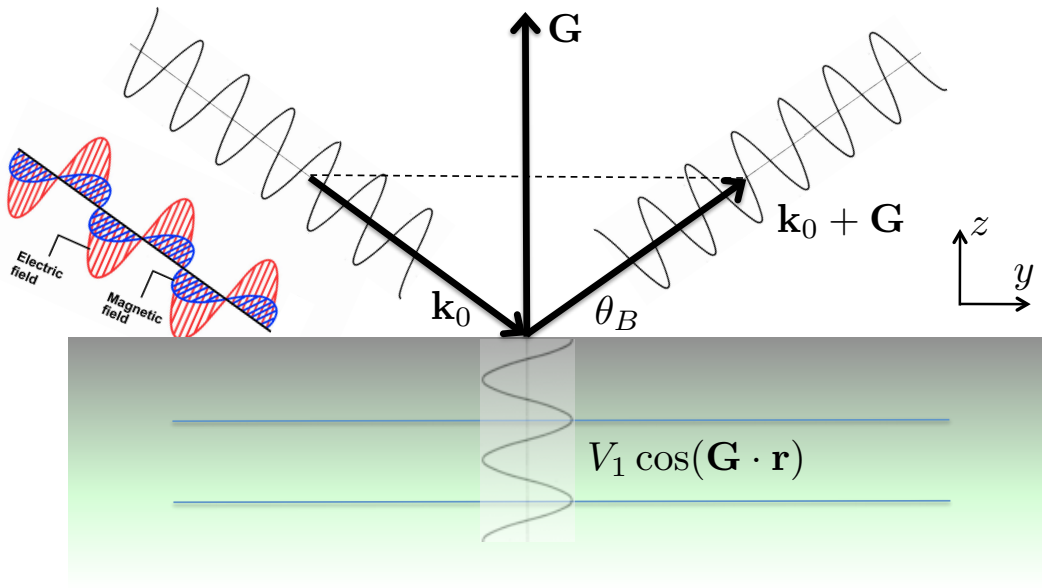


Fig. 1: *A setup of a simple diffraction experiment. An incoming wave with wave vector* $\mathbf{k}_0$ *and energy* $\hbar\omega_0$ *is scattered at a semi-infinite crystal, periodic along the* z*-direction. The wave on the left represents the case of X-ray radiation. In case of the Bragg scattering, the incoming wave is completely scattered into the reflected wave of the same energy and wave vector* $\mathbf{k}_0 + \mathbf{G}$*, where* $\mathbf{G}$ *is the reciprocal vector corresponding to the periodicity of the lattice.*

Substituting the Bloch wave into Eq. (1), we obtain a system of equations for the amplitudes A_n:

$$\left[\frac{\hbar^2}{2m}(\mathbf{k}+n\mathbf{G})^2 - \hbar\omega_0\right] A_n + \frac{V_1}{2}(A_{n-1} + A_{n+1}) = 0. \tag{3}$$

The system of equations above is completely identical to the system of equations for the coefficients of the wavefunction ψ expanded in the complete and orthonormal basis of quantum states $\{\psi_n\}$: $\psi = \sum_n A_n \psi_n$. Then, if the components of the Hamiltonian matrix $H_{nm} = \langle\psi_n|H|\psi_m\rangle$, wavefunction ψ solves the Schrödinger equation with energy E when the following system of equations is satisfied [13]:

$$(H_{nn} - E)A_n + \sum_{m\neq n} H_{nm}A_m = 0 \tag{4}$$

As follows from Eq. (3), in our case, the H_{nn} matrix elements are given by $\frac{\hbar^2}{2m}(\mathbf{k}+n\mathbf{G})^2 - \hbar\omega_0$, while the off-diagonal $H_{n,n\pm1}$ matrix elements assume the values of $V_1/2$.
In general, the system of equations Eqs. (3) and (4) has solutions only for certain values of E (or ω_0), if the Bloch vector $\mathbf{k}$ is given, which establishes the so-called dispersion relation $\omega_0(\mathbf{k})$ in a crystal. The wavefunction (2) with the vector $\mathbf{k}$ is the solution of the Schrödinger equation at energy E if the $\det(H_{nm} - E\delta_{nm}) = 0$. From this condition both $E(\mathbf{k}) = \hbar\omega_0(\mathbf{k})$ and the set of $\{A_n\}$ can be determined. The solution for $\{A_n\}$ for all n can be easily performed iteratively following Eq. (3) if, for example, A_0 and A_1 are pre-given. Effectively, the choice of A_0 and A_1 corresponds to the choice of the value of the wavefunction and its derivative when solving the second-order differential Schrödinger equation. It is known that generally, if arbitrary A_0 and

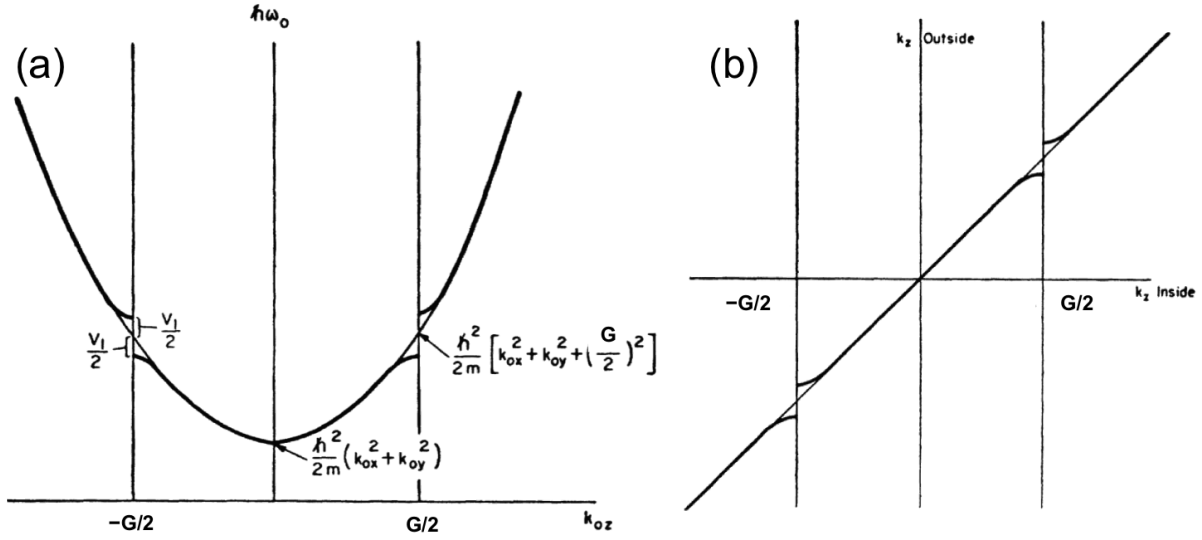


Fig. 2: *(a) Energy $\hbar\omega_0$ as a function of z component of the* **k***-vector for an electron wave inside the crystal. (b) The z component of the wave vector outside the crystal (k_{z0}) as a function of the z component of the wave vector inside the crystal (k_z). Taken from [9].*

A_1 are specified, the value of the wavefunction can be unbounded at infinity, however, upon an appropriate choice of ω_0 and A_1/A_0 ratio, the series (3) can be converged to provide a bounded wavefunction. Such frequencies ω_0 specify the energy spectrum of our system at $\mathbf{k}$.
Usually, in the theory of the *X-ray diffraction* and *electron diffraction* one of the main approximations which is made is the smallness of the V_1 (or its equivalent for X-rays, see last section). This allows to use approximations which are equivalent to the perturbation theory expressions. In the drastic case of $V_1 = 0$ (free space) the solution is trivial: $A_0 = 1$, $A_n = 0, n > 1$, and $\hbar^2\mathbf{k}^2/2m = \hbar\omega_0 = \hbar^2{\mathbf{k_0}}^2/2m$. When V_1 is sufficiently small we can derive approximate expression for the energy:

$$\hbar\omega_0 = \frac{\hbar^2\mathbf{k}^2}{2m} - \frac{V_1}{4}\left[\frac{1}{\frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})^2 - \hbar\omega_0} + \frac{1}{\frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{G})^2 - \hbar\omega_0}\right] \tag{5}$$

and the only surviving coefficients in the first order with respect to V_1 are:

$$A_{\pm 1} = -\frac{V_1}{2}A_0\frac{1}{\frac{\hbar^2}{2m}(\mathbf{k}\pm\mathbf{G})^2 - \hbar\omega_0} \tag{6}$$

While for most of the energies $A_0 \gg A_{\pm 1}$, in the particular case of the *Bragg scattering*, when $(\mathbf{k}\pm\mathbf{G})^2 = {\mathbf{k_0}}^2$, Eqs. (3) and (5) have no definite value. In this case, the degenerate perturbation theory should be used. Note that in our interpretation of the equations (3) the Bragg condition means the equality of the diagonal elements of the Hamiltonian matrix for $n = 0$ and $n = 1$. In this case, employing the degenerate perturbation theory we can show that at the point where the Bragg condition, $\mathbf{k} = \pm\mathbf{G}/2$, is met, exactly two energy solutions are possible:

$$\hbar\omega_0 = \frac{\hbar^2}{2m}\mathbf{k}^2 \pm \frac{V_1}{2}, \tag{7}$$

while there are two major participants in the Bloch wave: $A_0/A_1 = 1$. The solution $\hbar\omega_0$ as a function of $\mathbf{k}$ for the both considered above cases of the Bragg scattering and away from it, is

shown in Fig. 2(a). We have to remark that without assuming that there are only two waves in the crystal in the vicinity of the Bragg point (as we shall see in the following, it is called the *two-beam approximation*), we can recover analogously the opening of the gaps in the spectrum at points $\pm n\mathbf{G}/2$, with the magnitude of the gap proportional to V_1^n. Thus, if V_1 is small, then the two-beam approximation is justified.

Lets try to interpret now the appearance of the gap in the energy spectrum of a perfect periodic crystal (through V_1 and $\mathbf{G}$) from the point of view of a simple diffraction experiment, depicted in Fig. 1. In this experiment a wave of electrons (X-rays etc.) of a certain energy $\hbar\omega_0$ and wavelength $\mathbf{k}_0$, is sent toward the (possibly even finite) film, while the intensities of the reflected (or even transmitted to the other side of the film) waves are measured. Theoretical treatment of this experiment lies in finding the solutions to the Schrödinger equation in the vacuum and in the film, which can be very complicated owing to the dynamical nature of scattering. Two necessary conditions have to be satisfied in order to uniquely solve the problem: (i) the solution of the Schödinger equation in each part of space is sought at an energy $\hbar\omega_0$, and (ii) the wavefunction and its normal space derivative have to be continuous at the boundaries of the crystal.

In our simple case, the boundary conditions, owing to the fact that we have periodicity in the crystal only along the z-axis, lead to an observation that the in-plane components of the wave vector have to be continuous: $k_x = k_{x0}, k_y = k_{y0}$, where $\mathbf{k}_0 = (k_{0x}, k_{0y}, k_{0z})$ and the wave vector of the wave inside the crystal $\mathbf{k} = (k_x, k_y, k_z)$. The z-components of the both wave vectors do not have to be the same however, and they are indeed different in the vicinity of $\pm\mathbf{G}/2$. This discrepancy is obvious looking at the energy dispersion in Fig. 2(a), in which the energy of the electrons inside the crystal as a function of k_z is given with the thick line, while the thin line stands for the energy of electrons as a function of k_{0z} in the vacuum. As we can see, in the vicinity of $\pm\mathbf{G}/2$ the constant energy line will cross the two energy profiles at two different k-points. The correspondence between the two k-vectors at a constant energy is given in Fig. 2(b).

If the energy of an incoming wave lies in the gap of the energy solutions in the crystal, the reflection of the incoming wave will be very large. In this case the waves inside the crystal will be decaying with the distance from the surface of the film, since the Bloch waves cannot be matched to the wave outside the crystal. When the energy is in the middle of the gap, which occurs exactly at the Bragg angle of the incoming wave of θ_B, see Fig. 1, the reflection is complete. The range of the angles of the incident wave, θ, for which the transmission is unfavorable can be found easily:

$$\sin\theta = \left[\sin^2\theta_B \pm \frac{V_1}{2\hbar\omega_0}\right]^{1/2}. \tag{8}$$

On the other hand, when the experiment operates away from the Bragg condition, the matching of the incoming wave to the Bloch wave inside the crystal can be perfectly done, and, as a result, the amplitude of the reflected beam is very small.

3 Scattering of electrons in a crystal: general case

In this section we elaborate in more detail how the dynamical scattering can be studied in a way suggested in the previous section.

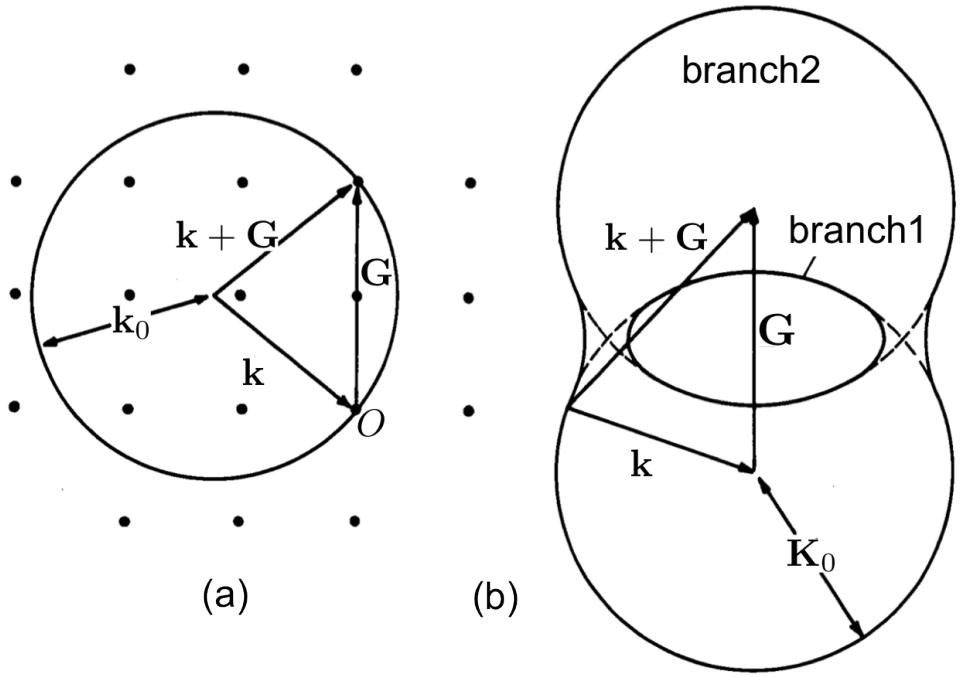


Fig. 3: *(a) Ewald sphere around the Bloch vector* $\mathbf{k}$ *of the radius corresponding to an incoming wave* $\mathbf{k}_0$*. The origin of the sphere is sometimes called the* Laue point*. (b) Dispersion surface for the two-beam case. Taken from [11].*

The motion of an electron in a solid is described by the Schrödinger equation Eq. (1):

$$H\psi = -\frac{\hbar^2}{2m}\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{9}$$

In an infinite periodic crystal the potential $V(\mathbf{r})$ is periodic with respect to translations by lattice vectors $\mathbf{R}$. The crystal potential can be split into contributions from different unit cells:

$$V(\mathbf{r}) = \sum_{\mathbf{R}} v(\mathbf{r} - \mathbf{R}) \tag{10}$$

where $v(\mathbf{r})$ is the potential of the Wigner-Seitz cell:

$$v(\mathbf{r}) = -\frac{Ze^2}{r} + \int_{V_C} d\mathbf{r}' \, \frac{e^2\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{11}$$

Here, the first term represents the attractive interaction with the nucleus of charge $Z = \int_{V_C} d\mathbf{r}' \, \rho(\mathbf{r}')$, and the second one is the repulsive interaction with the electron density ρ in the unit cell of the volume V_C. The crystal potential can be decomposed into Fourier series:

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} \tag{12}$$

where the set $\{\mathbf{G}\}$ is the reciprocal lattice. The expression for the $\mathbf{G}$-components of V reads:

$$V_{\mathbf{G}} = \frac{1}{V_C}\int_{V_C} e^{-i\mathbf{G}\mathbf{r}} V(\mathbf{r}) \, d\mathbf{r} = \frac{1}{V_C}\int_{\mathbb{R}^3} e^{-i\mathbf{G}\mathbf{r}} v(\mathbf{r}) \, d\mathbf{r} \tag{13}$$

Using expression Eq. (11) we come to:

$$V_{\mathbf{G}} = -\frac{4\pi e^2}{V_C}\frac{Z - f_{\mathbf{G}}}{G^2} \tag{14}$$

with

$$f_{\mathbf{G}} = \int_{V_C} e^{-i\mathbf{G}\mathbf{r}}\rho(\mathbf{r})\, d\mathbf{r} \tag{15}$$

The quantity $f_{\mathbf{G}}$ is known as *atomic scattering factor*, and it is normally smaller or equal than Z, ensuring that all the Fourier components of the potential are negative. For small $\mathbf{G}$'s the $V_{\mathbf{G}}$'s approach a constant, while for higher harmonics the interaction with the nucleus becomes increasingly important.
According to the Bloch's theorem valid in periodic crystals, the solution of the Schrödinger equation with a wave vector $\mathbf{k}$ can be represented as:

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} A_{\mathbf{G}}(\mathbf{k})e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}, \tag{16}$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a lattice-periodic function. The coefficients $A_{\mathbf{G}}(\mathbf{k})$ can be found by substituting the Bloch wave into the Schrödinger equation, leading thus to an infinite system of linear homogeneous equations:

$$\left\{E_{\mathbf{k}} - \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})\right\} A_{\mathbf{G}}(\mathbf{k}) = \sum_{\mathbf{G}'} V_{\mathbf{G}-\mathbf{G}'}A_{\mathbf{G}'}(\mathbf{k}) \tag{17}$$

By solving this system, for each $\mathbf{k}$ we can find a set of solutions $A_{\mathbf{G}}(\mathbf{k})$ which determine the Bloch wave, and the set of energies $E_{\mathbf{k}}$, with the latter one being commonly referred to as the *band structure*. This system has a solution only if the *dispersion equation* is satisfied:

$$\det\left[\left\{E_{\mathbf{k}} - \frac{\hbar^2}{2m}(\mathbf{k}+\mathbf{G})\right\}\delta_{\mathbf{G},\mathbf{G}'} - V_{\mathbf{G}-\mathbf{G}'}\right] = 0 \tag{18}$$

The effect of the *absorption* for electrons in crystals can be phenomenologically included by considering complex potential $V(\mathbf{r})$, complex $\mathbf{k}$-vectors and complex energies $E_{\mathbf{k}}$.
As already mentioned, for electron and X-ray diffraction the calculation of the Bloch waves and the band structures can be very much simplified since the energy is much larger than the potential, and, analogously to the previous section, we can apply the perturbation theory. Assuming first that $V(\mathbf{r}) = 0$, we get the following system of equations:

$$\left\{\mathbf{k}_0^2 - (\mathbf{k}+\mathbf{G})^2\right\} A_{\mathbf{G}} = 0, \tag{19}$$

where the energy $E_{\mathbf{k}}$ is given by an incoming wave with $\mathbf{k}_0$: $E_{\mathbf{k}} = \frac{\hbar^2}{2m}\mathbf{k}_0^2$. For given energy and given $\mathbf{k}$ the expression in brackets will in general not vanish for all $\mathbf{G}$ and there are no allowed waves. Nevertheless, for a certain $\mathbf{k}$ it may happen that $\mathbf{k}_0^2 = (\mathbf{k}+\mathbf{G})^2$ for one $\mathbf{G}$, for example $\mathbf{G} = 0$. Then $A_{\mathbf{G}} = \delta_{\mathbf{G},0}$ and the wave $e^{i\mathbf{k}\mathbf{r}}$ is allowed. For $V \neq 0$ the system of equations (18) comes to:

$$\left\{\mathbf{K}_0^2 - (\mathbf{k}+\mathbf{G})^2\right\} A_{\mathbf{G}} = \sum_{\mathbf{G}'\neq\mathbf{G}} v_{\mathbf{G}-\mathbf{G}'}A_{\mathbf{G}'}, \tag{20}$$

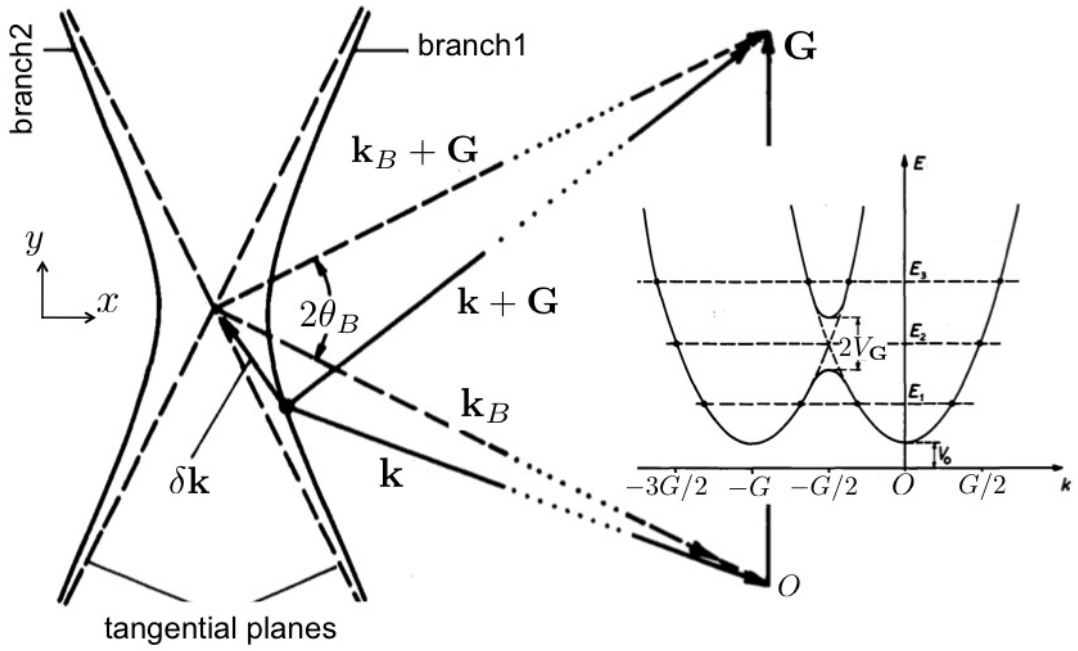


Fig. 4: *Left: Dispersion surface near the Bragg spot* $\mathbf{k}_B$*. Right: Bandstructure for the case on the right, for the direction of the Bloch vector along* y*. Taken from [11].*

with $v_\mathbf{G} = (2m/\hbar^2)V_\mathbf{G}$, and $\mathbf{K}_0^2 = \mathbf{k}_0^2 - v_0 = (2m/\hbar^2)(E_\mathbf{k} - V_0)$. It can be shown that for small $v_\mathbf{G}$ the secondary waves $A_\mathbf{G}$ for $\mathbf{G} \neq 0$ are small and we have only one strong beam $\psi = e^{i\mathbf{kr}}$ with a slightly renormalized $\mathbf{k}$-value.
However, if the condition $\mathbf{k}_0^2 \approx (\mathbf{k} + \mathbf{G})^2$ is fulfilled not only for the primary wave $\mathbf{G} = 0$ but also for other secondary waves $\mathbf{G} \neq 0$, then these waves may also become strong and the perturbation theory breaks down. This condition means that for certain $\mathbf{G}$ the vector $\mathbf{k} + \mathbf{G}$ lies near or at the so-called *Ewald sphere*, see Fig. 3. Because the energies $\frac{\hbar^2}{2m}\mathbf{k}^2$ and $\frac{\hbar^2}{2m}(\mathbf{k} + \mathbf{G})^2$ are close in this case, the degenerate perturbation theory has to be applied, which takes into account on an equal basis all the excited waves. Here, we restrict ourselves to the so-called *two-beam case*, for which only two waves are prominent in the crystal: at $\mathbf{k}$ and $\mathbf{k} + \mathbf{G}$. This approximation allows to obtain main conclusions of the dynamical scattering theory in a simple way, yet without loosing generality. We also used the two-beam approximation in the previous section. For two beams, the equations for the amplitudes of the waves read:

$$(\mathbf{K}_0^2 - \mathbf{k}^2)A_0 = v_{-\mathbf{G}}A_\mathbf{G}, \quad (\mathbf{K}_0^2 - (\mathbf{k} + \mathbf{G})^2)A_\mathbf{G} = v_\mathbf{G}A_0, \tag{21}$$

while the dispersion equation reads:

$$(\mathbf{K}_0^2 - \mathbf{k}^2) \cdot (\mathbf{K}_0^2 - (\mathbf{k} + \mathbf{G})^2) = v_\mathbf{G} \cdot v_{-\mathbf{G}}. \tag{22}$$

From the dispersion equation it is clear that in the absence of the crystal potential for a given by $\mathbf{K}_0^2$ energy the allowed $\mathbf{k}$-vectors lie on a *dispersion surface* consisting of two intersecting spheres centered around the origin, and $\mathbf{G}$, both with the radius $\mathbf{K}_0^2$, see Fig. 4. Non-zero $\mathbf{G}$-component of the potential, on the other hand, forces the intersection surface between the two spheres, at which the Bragg condition $\mathbf{k}^2 = (\mathbf{k} + \mathbf{G})^2$ is fulfilled, to split into two surfaces, of which the outer branch 2 completely surrounds the inner branch 1. While at the Bragg point

$\mathbf{k} = \mathbf{k}_B$, by setting $\mathbf{k} = \mathbf{k}_B + \delta\mathbf{k}$ around the intersection, we find:

$$4(\mathbf{k}_B \cdot \delta\mathbf{k})((\mathbf{k}_B + \mathbf{G}) \cdot \delta\mathbf{k}) = v_{\mathbf{G}} \cdot v_{-\mathbf{G}}, \tag{23}$$

if we neglect the 3rd and 4th orders in $\delta\mathbf{k}$. For convenience, we decompose the $\delta\mathbf{k}$ into x and y components, see Fig. 4. Using the condition that $K_0 \sin\theta_B = G/2$ we find:

$$\delta k_x^2 - \delta k_y^2 \tan^2\theta_B = \frac{|v_{\mathbf{G}}|^2}{4K_0^2 \cos^2\theta_B}. \tag{24}$$

In this approximation, the dispersion surfaces are hyperbolas, the asymptotes of which are the tangential planes of the spheres, see Fig. 4. The smallest separation between the two branches is

$$\Delta k = \frac{|v_{\mathbf{G}}|}{K_0 \cos\theta_B} \tag{25}$$

The distance $d_{ext} = \frac{2\pi}{\Delta k}$, over which the two Bloch waves from the opposite branches get a phase difference 2π, is called the *extinction length*.
From the dispersion equation we get the energy as a function of $\mathbf{k}$ as follows:

$$\frac{2m}{\hbar^2}E(\mathbf{k}) - v_0 = K_0^2 = \frac{1}{2}(\mathbf{k}^2 + (\mathbf{k}+\mathbf{G})^2 \pm \frac{1}{2}\sqrt{(\mathbf{k}^2 - (\mathbf{k}+\mathbf{G})^2)^2 + 4|v_{\mathbf{G}}|}, \tag{26}$$

showing that for $\mathbf{k}^2 = (\mathbf{k}+\mathbf{G})^2$ there is a band gap of the width $\Delta E = 2V_{\mathbf{G}}$, as in the previous section. We have plotted the bandstructure according to the relation above along the direction of $\mathbf{G}$ in Fig. 4. When the energy E_1 is below the gap, we get four allowed k-values. For the energy in the gap E_2 we obtain only two allowed k-values, while we restore the situation for E_1 when we go above the band gap (E_3).
From the previous formulas also the expressions for the amplitudes of the waves in the crystal can be derived:

$$A_0 = \frac{1}{\sqrt{2}}\sqrt{1 \mp \frac{W}{\sqrt{1+W^2}}}, \quad A_{\mathbf{G}} = \pm\mathrm{sign}(v_{\mathbf{G}})\frac{1}{\sqrt{2}}\sqrt{1 \pm \frac{W}{\sqrt{1+W^2}}}, \tag{27}$$

where an important parameter W, which measures the deviation from the Bragg angle, θ_B (see Fig. 4), is given by:

$$W = \frac{\delta\mathbf{k} \cdot \mathbf{G}}{|v_{\mathbf{G}}|}. \tag{28}$$

The limiting case of $|W| \to \infty$ corresponds to the situation away from the Bragg spot.
For the case of exact Bragg condition $W = 0$ the two Bloch waves are:

$$\psi^I(\mathbf{r}) = \sqrt{2}e^{i(\mathbf{k}+\frac{\mathbf{G}}{2})\mathbf{r}} \cdot \cos\left(\frac{\mathbf{G}\mathbf{r}}{2}\right), \quad \psi^{II}(\mathbf{r}) = -i\sqrt{2}e^{i(\mathbf{k}+\frac{\mathbf{G}}{2})\mathbf{r}} \cdot \sin\left(\frac{\mathbf{G}\mathbf{r}}{2}\right) \tag{29}$$

In case of electron diffraction, the Fourier components of the potential are negative, and the ψ^I will be positioned on the outer branch, while ψ^{II} lies at the inner branch. This situation will be reversed for X-ray diffraction, in which case $v_{\mathbf{G}}$ is positive. Characteristic for both branches are the cosine and sine modulation functions. Owing to these modulation functions, ψ^I is always maximal at the atomic positions of the reflecting planes (situated along $\mathbf{G}$), while it is small in between the planes, see Fig. 5. The situation is reversed for ψ^{II}, which is maximal between the

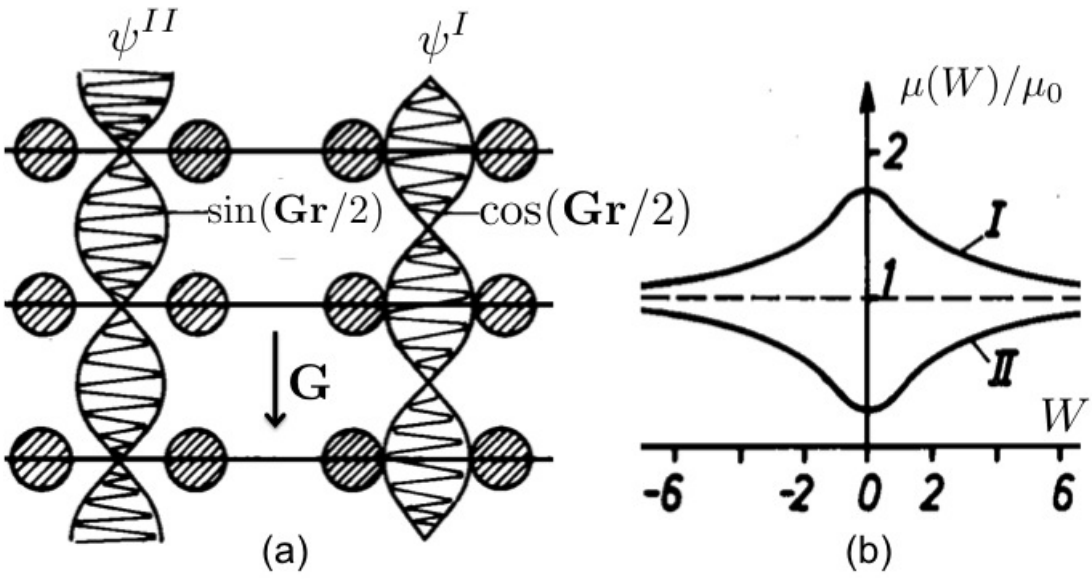


Fig. 5: *(a) Modulation factors of ψ^I and ψ^{II}. (b) Absorption for Bloch wave ψ^I and ψ^{II} at $\mu_{\mathtt{G}} = 3/4\mu_0$. Taken from [11].*

atomic planes, while it almost does not interact with atomic potential. Both waves, of course, have the same energy, however, the Bloch vector of ψ^I is larger than that of ψ^{II}, since the solutions are positioned on different branches. This manifests the larger kinetic energy of the first wave, which is compensated by larger but negative energy coming from interaction with the atoms.

For an absorbing crystal it is clear that the absorption will be very much different for ψ^I and ψ^{II}. The absorption of wave ψ^I will be higher than normal since an electron in the Bloch wave spends more time on the atoms, while wave ψ^{II} manages to avoid the atoms and has therefore a smaller than average absorption, resulting in a so-called *anomalous transmission* (or *anomalous absorption*).

4 Transmission and reflection from crystal slabs

In order to find the intensities of the reflected and transmitted waves, we have to rigorously consider the scattering on the incident wave by a finite crystal of volume Ω. The potential in the whole space in a plausible approximation can be written as $V(\mathbf{r}) = \theta_\Omega(\mathbf{r})V_\infty(\mathbf{r})$, where $\theta_\Omega(\mathbf{r})$ is the step function equal to 1 in the crystal and 0 outside of it, while $V_\infty(\mathbf{r})$ is the potential of the infinite crystal. Since the solutions of the Schrödinger equation in the vacuum and in the crystal are plane waves and Bloch waves, respectively, the wavefunction in the vacuum can be represented as:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}_0\mathbf{r}} + \sum_i R_i e^{i\mathbf{K}_i\mathbf{r}}, \tag{30}$$

while in the crystal

$$\psi(\mathbf{r}) = \sum_j P_j \psi_{\mathbf{k}_j}(\mathbf{r}), \tag{31}$$

with the energy $E = \frac{\hbar^2}{2m}\mathbf{k}_0^2 = \frac{\hbar^2}{2m}\mathbf{K}_i^2 = E(\mathbf{k}_j)$. How many waves in the vacuum and in the crystal are excited by an incident wave depends very much on the shape of the crystal, and the coefficients R_i and P_j have to be determined via imposing the boundary conditions. Namely, the wavefunction has to be continuous across the surface between the vacuum and the crystal, $\psi_{vac}(\mathbf{r}) = \psi_{crys}(\mathbf{r})$, and the same applies to the current through the surface, $\mathbf{n}(\mathbf{r})\frac{d\psi_{vac}(\mathbf{r})}{d\mathbf{r}} = \mathbf{n}(\mathbf{r})\frac{d\psi_{crys}(\mathbf{r})}{d\mathbf{r}}$, where $\mathbf{n}(\mathbf{r})$ is the surface normal. In order to take the evanescent states into consideration, we have to rely on the fact that the wave vectors in the vacuum and in the crystal can be in general complex.

The solution of the problem can be found by employing the integral equation for $\psi(\mathbf{r})$ in terms of the Green's function:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}_0\mathbf{r}} + \int_{\mathbb{R}^3} d\mathbf{r}'\, G_0(\mathbf{r}-\mathbf{r}')V(\mathbf{r}')\psi(\mathbf{r}), \tag{32}$$

where $G_0(\mathbf{r}-\mathbf{r}')$ is the Green's function given by:

$$G_0(\mathbf{r}-\mathbf{r}') = -\frac{2m}{\hbar^2}\frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}. \tag{33}$$

Owing to the specific shape of our potential, as can be seen from (32) the wave in the vacuum consists of the incident wave and the sum of the spherical waves with the energy E, outgoing from the crystal. On the other hand, in the crystal, the spherical waves completely suppress the $e^{i\mathbf{k}_0\mathbf{r}}$ wave (so-called extinction theorem), while forming at the same time the Bloch solutions [11].

Equation (32) allows for a simple interpretation. The incident plane wave $e^{i\mathbf{k}_0\mathbf{r}}$ is compensated by spherical waves emitted from scattering at position $\mathbf{r}'$ in the crystal. The intensity of these scattered waves is proportional to the interaction potential $V(\mathbf{r}')$ and the amplitude of the total wave field $\psi(\mathbf{r}')$ at that point. Note that the wavefunction ψ appears both on the right and left hand side of Eq. (32). In *kinematical theory* of diffraction, the ψ on the right hand side of Eq. (32) is replaced by the incident wave $e^{i\mathbf{k}_0\mathbf{r}}$ itself, which corresponds to so-called ***first Born approximation***. The physics behind the 1^{st} Born approximation is the assumption that the incoming wave scatters only *once* on the crystal potential before forming the scattered wave ψ^1. In most of the cases, especially for neutron scattering, the kinematic approximation works very well. On the other hand, for X-ray and especially electron scattering the kinematic theory very often does not provide good results. In this case using the theory which goes beyond the 1^{st} Born approximation, namely, *dynamical scattering theory*, is necessary. This can be done, for example, by iteratively using the solution of the $(n-1)^{st}$ Born approximation in order to construct the solution of the next interaction via solving Eq. (32). This procedure corresponds to converging the so-called ***Born series***. For more details, see the manuscript of Prof. Stefan Blügel in this book [1].

Let us now consider the case of a finite crystal slab, which fills the space between $0 \leq z \leq d$, see Fig. 6 for setup. The potential in the crystal is thus periodic in the $x-y$ plane, with the reciprocal two-dimensional lattice $\{\mathtt{G}\}$. Correspondingly, each wavefunction in the crystal can be identified with a certain in-plane Bloch vector $\mathtt{k}$, and can be expanded in $\mathbb{R}^3$ as ($\mathtt{r}$ is the coordinate in-plane):

$$\psi_{\mathtt{k}}(\mathtt{r}, z) = \sum_{\mathtt{G}} \Gamma_{\mathtt{G}}(z) e^{i(\mathtt{k}+\mathtt{G})\mathtt{r}} \tag{34}$$

Owing to the boundary conditions, the Bloch vector is determined by the in-plane component of the incident wave $\mathbf{k}_0 = (\mathtt{k}, k_0^z)$ By substituting this equation into the Schrödinger equation

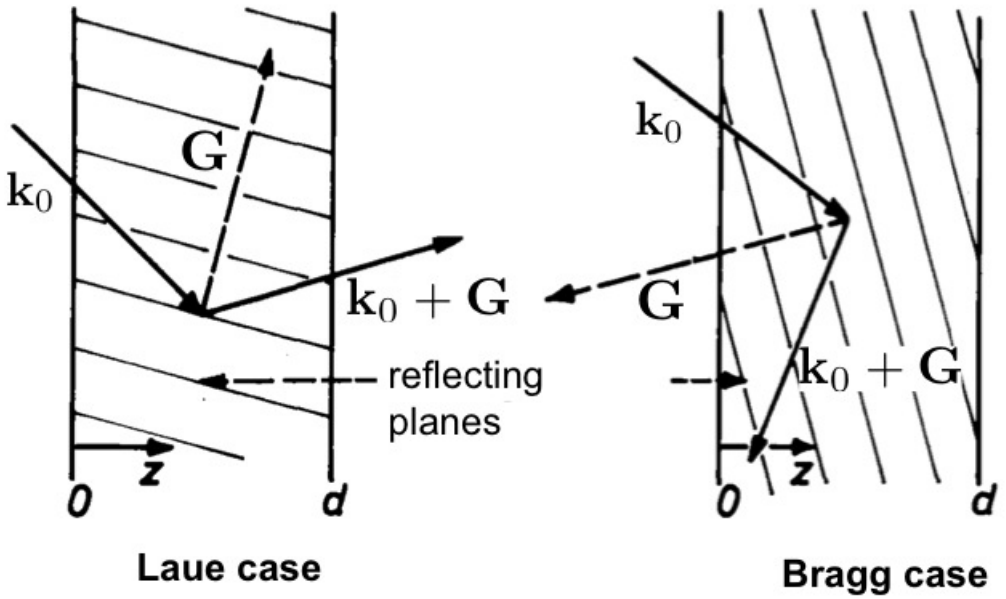


Fig. 6: *Incident wave* $\mathbf{k}_0$ *and reflected wave* $\mathbf{k}_0 + \mathbf{G}$ *for the Laue and Bragg case. Taken from [11].*

we obtain a set of equations for $\Gamma_{\mathtt{G}}(z)$:

$$\left[\partial_z^2 + (k_0^2 - (\mathtt{k} + \mathtt{G})^2)\right] \Gamma_{\mathtt{G}}(z) = \sum_{\mathtt{G}'} v_{\mathtt{G}-\mathtt{G}'}(z)\Gamma_{\mathtt{G}'}(z). \tag{35}$$

From this equation we can conclude that for each $\mathtt{G}$ the energy is split up into the energy for motion in-plane, $\frac{\hbar^2}{2m}(\mathtt{k}+\mathtt{G})^2$, and the remaining of the energy $\frac{\hbar^2}{2m}k_0^2$ for motion in the z-direction. Since $\mathtt{G}$ can be arbitrarily large, the energy along z can be negative, which corresponds to the bound modes decaying into the vacuum.
Equation (32) can be used to find the wavefunctions, and written for $\Gamma_{\mathtt{G}}(z)$ functions it reads:

$$\Gamma_{\mathtt{G}}(z) = \delta_{\mathtt{G},0}\, e^{izk_0^z} + \int dz'\, \frac{e^{iK_{\mathtt{G}}|z-z'|}}{2iK_{\mathtt{G}}} \sum_{\mathtt{G}'} v_{\mathtt{G}-\mathtt{G}'}(z')\Gamma_{\mathtt{G}'}(z'), \tag{36}$$

where

$$K_{\mathtt{G}} = \sqrt{k_0^2 - (\mathtt{k}+\mathtt{G})^2},\ \text{if } K_{\mathtt{G}}^2 > 0 \qquad K_{\mathtt{G}} = i\kappa_{\mathtt{G}},\ \kappa_{\mathtt{G}} > 0,\ \text{if } K_{\mathtt{G}}^2 < 0 \tag{37}$$

In the vacuum for $z < 0$ we get

$$\Gamma_{\mathtt{G}}(z) = \delta_{\mathtt{G},0}\, e^{izk_0^z} + R_{\mathtt{G}}\, e^{-izK_{\mathtt{G}}} \tag{38}$$

with

$$R_{\mathtt{G}} = \frac{1}{2iK_{\mathtt{G}}} \int_0^d dz'\, e^{+iK_{\mathtt{G}}z'} \sum_{\mathtt{G}'} v_{\mathtt{G}-\mathtt{G}'}(z')\Gamma_{\mathtt{G}'}(z'), \tag{39}$$

and analogously for $z > d$:

$$\Gamma_{\mathtt{G}}(z) = \delta_{\mathtt{G},0}\, e^{izk_0^z} + T_{\mathtt{G}}\, e^{+izK_{\mathtt{G}}} \tag{40}$$

with

$$T_{\mathtt{G}} = \frac{1}{2iK_{\mathtt{G}}} \int_0^d dz'\, e^{-iK_{\mathtt{G}}z'} \sum_{\mathtt{G}'} v_{\mathtt{G}-\mathtt{G}'}(z')\Gamma_{\mathtt{G}'}(z'). \tag{41}$$

Coefficients $R_{\mathtt{G}}$ and $T_{\mathtt{G}}$ are called the *reflection* and *transmission coefficients*, respectively. Only in the case when $K_{\mathtt{G}}^2 = k_0^2 - (\mathtt{k}+\mathtt{G})^2 > 0$ do we get a reflected plane wave with $\mathbf{K}_{\mathtt{G}}^- = (\mathtt{k}+\mathtt{G}, -K_{\mathtt{G}})$ and a transmitted plane wave $\mathbf{K}_{\mathtt{G}}^+ = (\mathtt{k}+\mathtt{G}, +K_{\mathtt{G}})$. The waves with complex $K_{\mathtt{G}}$ decrease exponentially into the vacuum. It is clear, that whereas we have only a finite number of reflected and transmitted waves, the number of decaying waves is infinite. A graphical construction of the waves $\mathbf{K}_{\mathtt{G}}^{\pm}$ is shown in Fig. 8, in which all vectors $\mathbf{K}_{\mathtt{G}}^{\pm}$ lie on the sphere of radius k_0. Assuming that the vector $\mathtt{G}$ lies in the plane of $\mathbf{k}_0$ and surface normal, the only four allowed real vectors $\mathbf{K}_{\mathtt{G}}^{\pm}$ and $\mathbf{K}_{2\mathtt{G}}^{\pm}$ are shown.

The kinematical scattering theory is constructed based on the substitution of $e^{iz'k_0^z}$ instead of $\Gamma_{\mathtt{G}'}(z')$ in the right hand side of Eq. (36). The intensities of the waves obtained within the kinematical theory can very much deviate from the predictions of the dynamical theory, especially for the case of Bragg reflection $k_0^z \pm K_{\mathtt{G}} + g = 0$, where g is the projection of $\mathbf{G}$ on the z-axis, see Fig. 6.

Inside the crystal, only those Bloch waves $\mathbf{k}_j$ with energy $E(\mathbf{k}_j) = \frac{\hbar^2}{2m}k_0^2$, are allowed. Moreover the in-plane component of the Bloch vector can differ from $\mathtt{k}$ only by an in-plane reciprocal vector $\mathtt{G}$. Since the energy of the Bloch states are periodic in reciprocal space, we can assume that $\mathbf{k}_j = (\mathtt{k}, k_{jz})$, where k_{jz} are determined from the energy conservation. The P_j coefficients in Eq. (31) can be also determined by matching the waves inside and outside the crystal at both vacuum boundaries.

4.1 Symmetric Laue case

Let us consider now the case when only one Bragg reflection $\mathbf{G}$ is excited, see Fig. 6. In this case we can apply the two-beam approximation. In principle, for an incident from $-\infty$ wave there are two situations to consider: when the crystal-reflected $\mathbf{k}_0 + \mathbf{G}$ wave propagates forward (*Laue case*) or backwards (*Bragg case*), see Fig. 6. In this subsection we discuss the so-called symmetric Laue case, for which $\mathbf{G}$ lies in-plane, and can be thus represented as $\mathbf{G} = (\mathtt{G}, 0)$. In this case the reflecting atomic planes are perpendicular to the surface of the slab.

For symmetric Laue case the setup of the problem is presented in Fig. 7 in the reciprocal space. We have the following waves in the vacuum. For $z > d$: the transmitted wave $\mathbf{k}_0 = \mathbf{K}_0^+ = (\mathtt{k}, k_0^z)$ (given by vector AO) and the Bragg reflected wave $\mathbf{K}_{\mathtt{G}}^+ = (\mathtt{k}+\mathtt{G}, K_{\mathtt{G}})$ (given by vector BH). For $z < 0$ we have the incident wave $\mathbf{k}_0$, the surface reflected wave $\mathbf{K}_0^- = (\mathtt{k}, -k_0^z)$ (given by vector DO) and the surface reflected wave of the forward propagating $\mathbf{K}_{\mathtt{G}}^+$ wave, $\mathbf{K}_{\mathtt{G}}^- = (\mathtt{k}+\mathtt{G}, -K_{\mathtt{G}})$ (given by vector CH). In the crystal, we get four Bloch waves $\mathbf{k}_j = (\mathtt{k}, k_{jz})$, where the k_{jz} are determined from the dispersion relation:

$$(K_0^2 - \mathtt{k}^2 - k_{jz}^2)(K_0^2 - (\mathtt{k}+\mathtt{G})^2 - k_{jz}^2) = |v_{\mathtt{G}}|^2. \tag{42}$$

The dispersion surface in the vacuum (thin lines) and in the crystal (thick lines) is shown together with vectors $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4$ (numbers) in Fig. 7. Characteristically for electrons the radius $K_0 = \sqrt{k_0^2 - v_0}$ is larger than k_0 (opposite for X-rays). However, the difference $(K_0 - k_0)/k_0 \approx -v_0/(2k_0^2)$ is extremely small or high energy electrons. For example, if we have $E = 100$ eV and $-v_0 = 10$ eV one gets for the latter ratio a value of $5 \cdot 10^{-5}$. Therefore vectors $\mathbf{k}_0, \mathbf{k}_1, \mathbf{k}_2$ in Fig. 7 are basically the same.

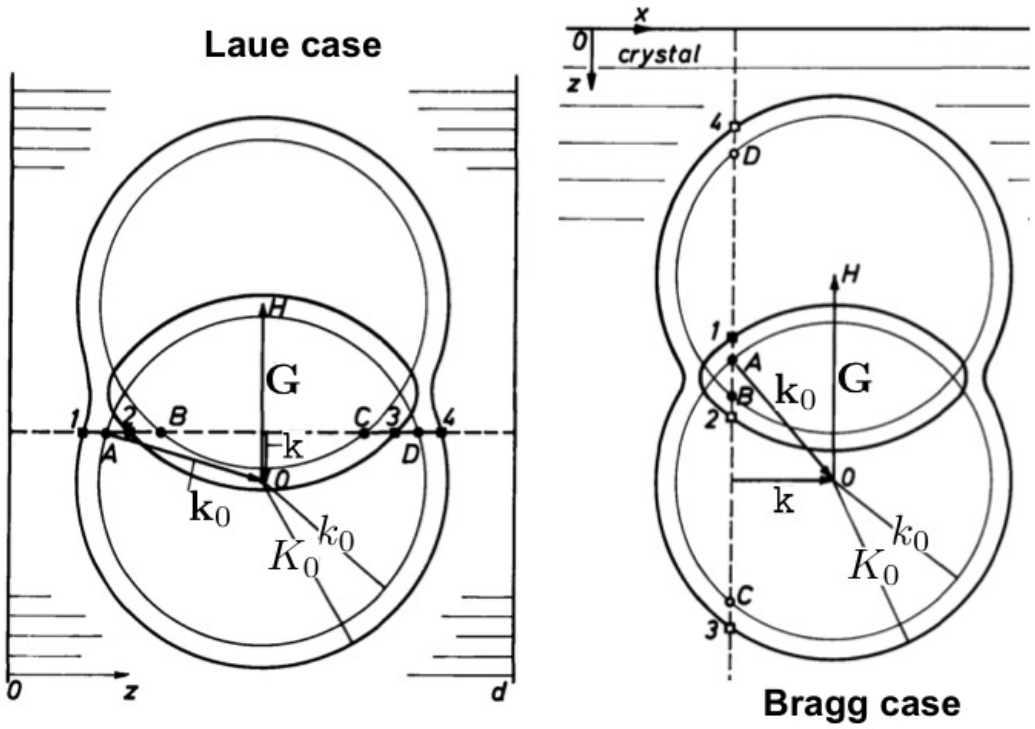


Fig. 7: *Dispersion surface in the crystal and in the vacuum for symmetric Laue case and symmetric Bragg case. Thin lines stand for the dispersion surface in the crystal, while thick lines stand for the dispersion surface in the crystal. Taken from [11].*

It can be shown by solving equations for transmission, reflection and R_j coefficients from the beginning of this section, that for the symmetric Laue case with a high degree of accuracy there are only two strong waves in the vacuum: $\mathbf{k}_0$ and $\mathbf{K}_\mathtt{G}^+$, while in the crystal the only two Bloch waves which can be considered are $\mathbf{k}_1$ and $\mathbf{k}_2$ (see Fig. 8(b)). The intensity of the transmitted waves can be calculated:

$$|T_0|^2 = \frac{W^2}{1+W^2} + \frac{1}{1+W^2}\cos^2\left(\frac{d\Delta k\sqrt{1+W^2}}{2}\right), \tag{43}$$

and

$$|T_\mathtt{G}|^2 = \frac{1}{1+W^2}\sin^2\left(\frac{d\Delta k\sqrt{1+W^2}}{2}\right), \tag{44}$$

where W and Δk are given by Eqs. (28) and (25). Exactly at the Bragg condition $W = 0$

$$|T_0|^2 = \cos^2\left(\frac{d\Delta k}{2}\right), \text{ and } \quad |T_\mathtt{G}|^2 = \sin^2\left(\frac{d\Delta k}{2}\right). \tag{45}$$

These are the so-called *Pendellösungen* (pendulum solutions), manifesting the exchange of intensity between the two transmitted beams as a function of the film thickness d, see Fig. 8. The period of this oscillation is the extinction length d_{ext}. By writing the wave field inside the crystal for the exact Bragg condition $\mathbf{k}_B = (\mathbf{k}_1 + \mathbf{k}_2)/2$:

$$\psi(\mathbf{r}) = e^{i\mathbf{k}_B\mathbf{r}}\cos\left(\frac{z\Delta k}{2}\right) + ie^{i(\mathbf{k}_B+\mathbf{G})\mathbf{r}}\sin\left(\frac{z\Delta k}{2}\right), \tag{46}$$

we observe again the pendulum solution with the thickness-dependent amplitude of the both waves which constitute the crystal wave field, see Fig. 8.

Effect of absorption. In case of a crystal with absorption it is important to realize the the Bloch waves $\mathbf{k}_1$ and $\mathbf{k}_2$ are absorbed differently, since they lie on different branches of the dispersion surface, as discussed in the previous section. It can be shown that as a function of W the absorption coefficient which describes the exponential decay of the intensity of the beam for both solutions is

$$\mu^{I,II}(W) = \mu_0 \pm \frac{1}{\sqrt{1+W^2}}\mu_{\mathtt{G}}, \tag{47}$$

where μ_0 and $\mu_{\mathtt{G}}$ are proportional to the corresponding parts of the Fourier components $\tilde{v}_0$ and $\tilde{v}_{\mathtt{G}}$ of the complex crystal potential which describes the absorption process. Clearly, μ_0 corresponds to the average strength of the absorption of a single plane wave by an "averaged" potential. An example of a behavior of the ratio $\mu^{I,II}(W)/\mu_0$ is given in Fig. 5(b). We can see that the difference in the absorption of the solutions on both branches is particularly strong in the vicinity of the Bragg condition. If the thickness of the crystal is rather large the wave lying on the 1^{st} branch will be therefore almost completely absorbed. Since the pendulum solutions arise due to the interference of the waves on both branches, in the limit of a semi-infinite crystal the oscillations in the intensity of two transmitted beams are diminished.

In greater detail, if we consider the case of exact Bragg condition, $W = 0$, the intensities of the transmitted waves as a function of the thickness of the crystal d can be decomposed into "non-oscillating" (no) and "oscillating" (o) parts, which can be calculated to be:

$$|T_0|^2_{no} \sim e^{-(\mu_0+\mu_{\mathtt{G}})d} + e^{-(\mu_0-\mu_{\mathtt{G}})d}, \quad |T_0|^2_{o} \sim e^{-\mu_0 d}\cdot\cos^2\left(\frac{d\Delta k}{2}\right), \tag{48}$$

$$|T_{\mathtt{G}}|^2_{no} \sim e^{-(\mu_0+\mu_{\mathtt{G}})d} + e^{-(\mu_0-\mu_{\mathtt{G}})d}, \quad |T_{\mathtt{G}}|^2_{o} \sim e^{-\mu_0 d}\cdot\sin^2\left(\frac{d\Delta k}{2}\right). \tag{49}$$

Typical measured curves for the intensities look therefore similar to the case of X-ray scattering in silicon, see Fig. 10(a). What we observe from the above expressions is that on the non-oscillating background intensity which decays as $e^{-(\mu_0+\mu_{\mathtt{G}})d} + e^{-(\mu_0-\mu_{\mathtt{G}})d}$ with d, there is an oscillating term which decays proportionally to $e^{-\mu_0 d}$. It is important to realize that, away from the Bragg condition there is no interference of waves I and II, which leads to different absorption and oscillatory term in above expressions. Thus, away from the Bragg spot, the intensity of the transmitted waves is simply proportional to $e^{-\mu_0 d}$, owing to the "smearing" of the wave inside the crystal over the unit cell.

Imagine now a situation in which the angle of incidence is varied and the intensity of the transmitted beam (for certainty lets assume it is given by $|T_{\mathtt{G}}|^2$) is measured for a thick crystal with thickness d. Away from θ_B this intensity is proportional to $e^{-\mu_0 d}$, but once the angle of incidence approaches θ_B, the dominant contribution to intensity becomes proportional to $e^{-(\mu_0-\mu_{\mathtt{G}})d}$ (lets say $0 < \mu_{\mathtt{G}} < \mu_0$). Thus, experimentally, we see a peak in the measured intensity at the Bragg angle. This constitutes the essence of the *Borrmann effect*, observed first for X-rays. For X-rays, the absorption process in highly localized since it comes mostly from the excitations of electrons in inner atomic shells. Thus, the Fourier coefficients of the absorption function decay very slowly, and the value of $\mu_{\mathtt{G}}$ can be very close to μ_0, leading to a very pronounced peak in the measured intensity of the transmitted wave. For electrons the effect is more modest, since the absorption comes mostly from plasmon excitations.

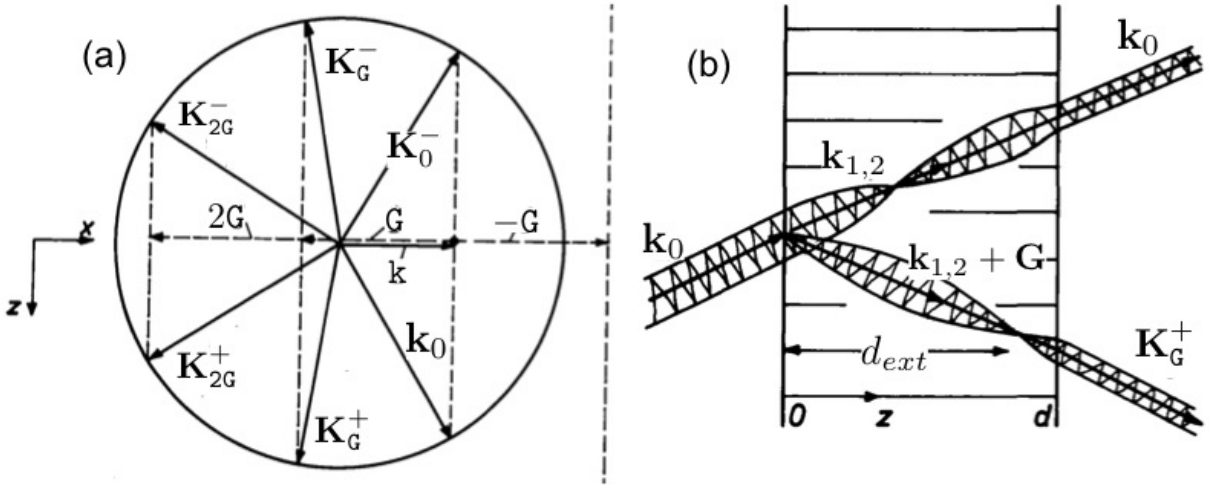


Fig. 8: *(a) Plane waves $\mathbf{K}_G^{\pm}$ in the vacuum. (b) Wave fields in the symmetrical Laue case at exact Bragg condition ($W = 0$). Taken from [11].*

4.2 Symmetric Bragg case

In symmetric Bragg case the reflecting planes are parallel to the surface of the slab, i.e. $\mathbf{G} = (0, 0, -G)$, see Fig. 7. Here for simplicity we consider only the case of a semi-infinite crystal, which is exactly the case we considered in the first section.

The incident wave in the vacuum $\mathbf{k}_0 = (\mathtt{k}, k_0^z)$ is shown with vector AO in Fig. 7. All other allowed waves in the vacuum have to lie on one of the two spheres with radius k_0 and have the same in-plane component $\mathtt{k}$. The wave with vector DH is not a new wave since it is the same vector as AO. Since $BH = CO$ we have only one new wave in the vacuum, $\mathbf{K}_0^- = (\mathtt{k}, -k_0^z) = BH$. In the crystal we can in principle get four Bloch waves $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$ and $\mathbf{k}_4$, marked with corresponding numbers in Fig. 7. However, $\mathbf{k}_3$ and $\mathbf{k}_4$ do not lie in the first Brillouin zone, thus, they do not have to be considered. Moreover, it can be shown that only $\mathbf{k}_1$ has a positive group velocity and it is the only wave which should be considered in the crystal. In a slab of finite thickness both Bloch waves are allowed, leading to oscillations in the reflected intensity, analogously to the Laue case.

Upon decreasing the angle of incidence, or, equivalently, increasing the in-plane component $\mathtt{k}$ the vector $\mathbf{k}_1$ moves along the dispersion surface towards the neck between the two branches. When further increasing $\mathtt{k}$ the line of constant $\mathtt{k}$ enters the gap between the two branches before crossing the outer branch at the neck. Without absorption, the reflection coefficient can be calculated to be:

$$|R_0|^2 = \left| \frac{1}{-y \pm \sqrt{y^2 - 1}} \right|^2 , \tag{50}$$

where

$$y = 2\frac{\delta k_x}{\Delta k} = \frac{2K_0 \cdot \cos\theta_B \cdot \delta k_x}{|v_{\mathbf{G}}|} \tag{51}$$

according to Eqs. (24) and (25). In the last equation sign "$-$" refers to the inner branch while "$+$" stands for the outer branch and the gap. For $|y| < 1$, i.e. inside the gap, the reflection coefficient $|R_0|^2 = 1$, while when the line $\mathtt{k} = \mathrm{const}$ deviates from the gap, the reflection coefficient decays very rapidly, see Fig. 9(a). The shape of R_0 as a function of the angle of

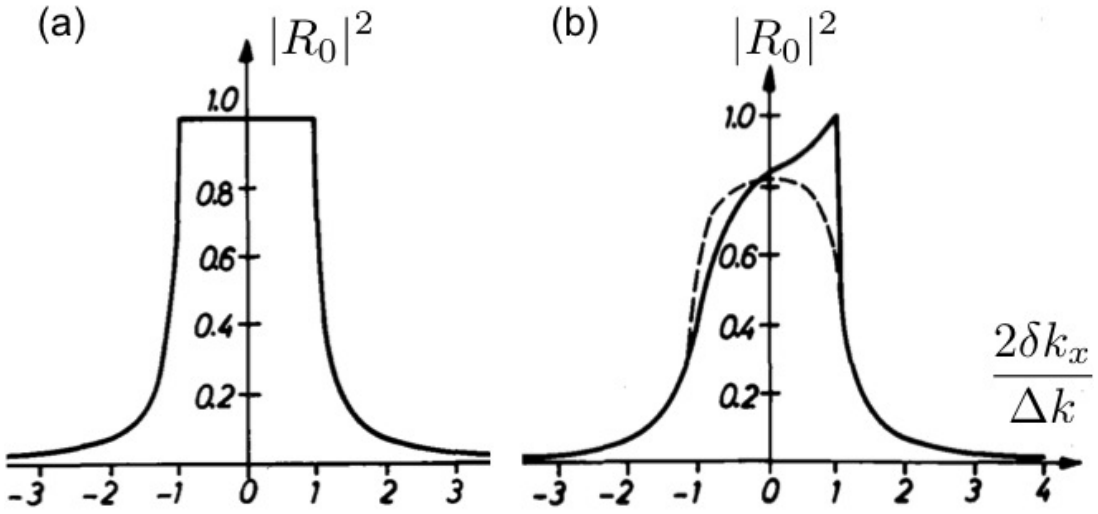


Fig. 9: *(a) Reflection coefficient $|R_0|^2$ in symmetric Bragg case without absorption. (b) Reflection coefficient $|R_0|^2$ in symmetric Bragg case with absorption with $\tilde{v}_0 = \tilde{v}_{\mathbf{G}}$ (solid line) and $\tilde{v}_{\mathbf{G}} = 0$ (dashed line). Taken from [11].*

incidence is known as the *Darwin profile*. The width of the gap where $|R_0|^2 = 1$, which is also called the *Darwin width*, exactly corresponds to the range of angle of incidence we derived in the first section, in which the perfect reflection is observed.
In case with absorption, it can be shown, that absorption is especially effective at the edges of the two branches, as can be seen from Fig. 9(b). Overall, the deviation of $|R_0|^2$ from one is inversely proportional to $\sqrt{v'_{\mathbf{G}}}$, where $v_{\mathbf{G}} = v'_{\mathbf{G}} + i\tilde{v}_{\mathbf{G}}$ is the Fourier component of the complex crystal potential. It is clear that while the absorbing potential is uniform, that is, it is given only by $\tilde{v}_0$, the waves on both dispersion branches are absorbed equivalently, as can be seen in Fig. 9(b) (dashed line). On the other hand, when the absorbing potential assumes the crystal structure, e.g. for $\tilde{v}_{\mathbf{G}} = \tilde{v}_0$, the waves I and II which have different distribution with respect to positions of the atoms, are absorbed differently, and an asymmetry in the reflection coefficient as a function of δk_x, or, equivalently, the angle of incidence, is evident, see Fig. 9(b) (solid line). An experimental example of such behavior can be clearly seen for instance in case of X-ray scattering in GaAs, for which the *rocking curve* (reflected intensity as a function of the angle of incidence) is shown in Fig. 10(b).
Physically, perfect reflection as that one in Fig. 9(a) is observed for neutrons scattered by thick slabs, since for neutrons the absorption is negligible. For X-rays the photoelectric absorption is concentrated mainly at the inner dispersion branch which results in anomalous transmission effect, and asymmetrical curves of Fig. 9(b) can be seen experimentally.

5 X-ray scattering

The theory of X-ray diffraction is quite analogous to the theory of electron diffraction with the exception that the wave fields are vector fields in nature (see Fig. 1). The interaction of X-rays with the medium arises via excitation of atomic electrons. The frequencies of motion of electrons are of the order of $\omega_0 \approx v_e/a_B$, where v_e is the velocity of electron motion around

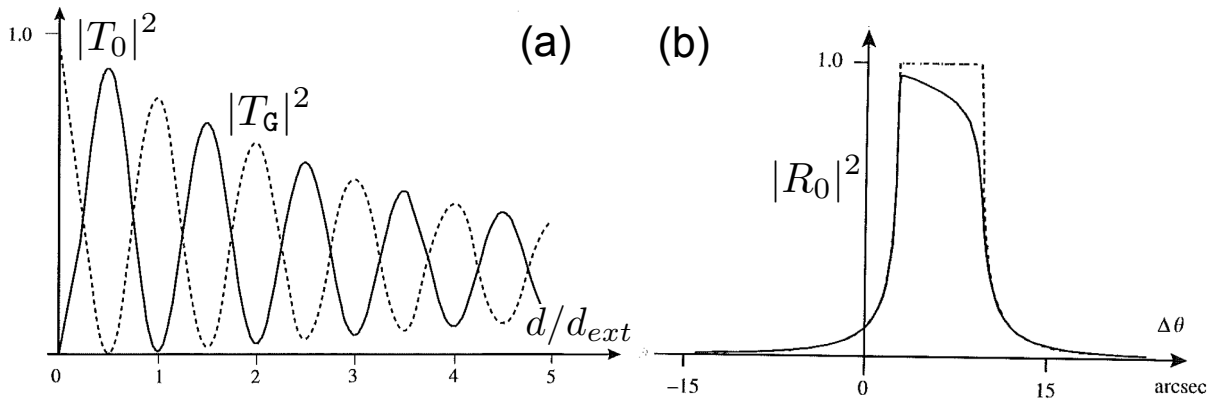


Fig. 10: *(a) Intensities of transmitted beams as a function of crystal thickness for 220 reflection of silicon with CuK_α radiation, $d_{ext} = 15.26$ μm. (b) Rocking curve (reflected intensity) for a thick crystal of 400 reflection of GaAs with CuK_α radiation (solid curve) compared to a theoretical calculation without absorption. Darwin width is 7.54 arcsec. Taken from [8].*

the atom, and a_B is Bohr's radius. The wavelength of X-rays is comparable to a_B while their frequencies are of the order of $2\pi c/\lambda$ with c as the speed of light, that is, much higher than those of the electrons. Thus, electrons can be treated as free. This simplification leads to the following formulation of the Maxwell's equations, which have to be solved instead of the Schrödinger equation for electrons (assuming harmonic time-dependence of the X-ray field with the frequency ω):

$$\nabla \times \mathbf{E} = i\frac{\omega}{c}\mathbf{H}, \quad \nabla \times \mathbf{H} = -i\frac{\omega}{c}\mathbf{D}, \tag{52}$$

where the electric displacement field $\mathbf{D} = \varepsilon\mathbf{E}$ and ε is the dielectric function of the medium:

$$\varepsilon(\mathbf{r}, \omega) = 1 - \frac{4\pi e^2}{m\omega^2}\rho(\mathbf{r}), \tag{53}$$

with $\rho(\mathbf{r})$ as the electronic density. Additionally,

$$\nabla \cdot \mathbf{D} = \nabla \cdot \mathbf{H} = 0, \quad \nabla \times \nabla \times \mathbf{E} = (\omega/c)^2\mathbf{D}. \tag{54}$$

Deviation of ε from unity is actually very small for X-rays and the quantity $\chi = \varepsilon - 1$ normally varies between 10^{-6} and 10^{-4}. Thus, we can write that

$$\mathbf{E} = \frac{1}{\varepsilon}\mathbf{D} \approx \mathbf{D} - \chi\mathbf{D}. \tag{55}$$

This equation is a great simplification, which, substituted into previous equations leads to an equation for $\mathbf{D}$ alone:

$$(\nabla^2 + k_0^2)\mathbf{D}(\mathbf{r}) = -\nabla \times \nabla \times [\chi(\mathbf{r})\mathbf{D}(\mathbf{r})] \tag{56}$$

In an infinite periodic crystal the charge density as well as the dielectric function are periodic and can be expanded in terms of the Fourier series in analogy to the potential for electrons Eq. (12). Analogously to electrons, the solution to Eq. (56) can be sought in the shape of a Bloch wave:

$$\mathbf{D}_\mathbf{k}(\mathbf{r}) = \sum_\mathbf{G} \mathbf{D}_\mathbf{G} e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}. \tag{57}$$

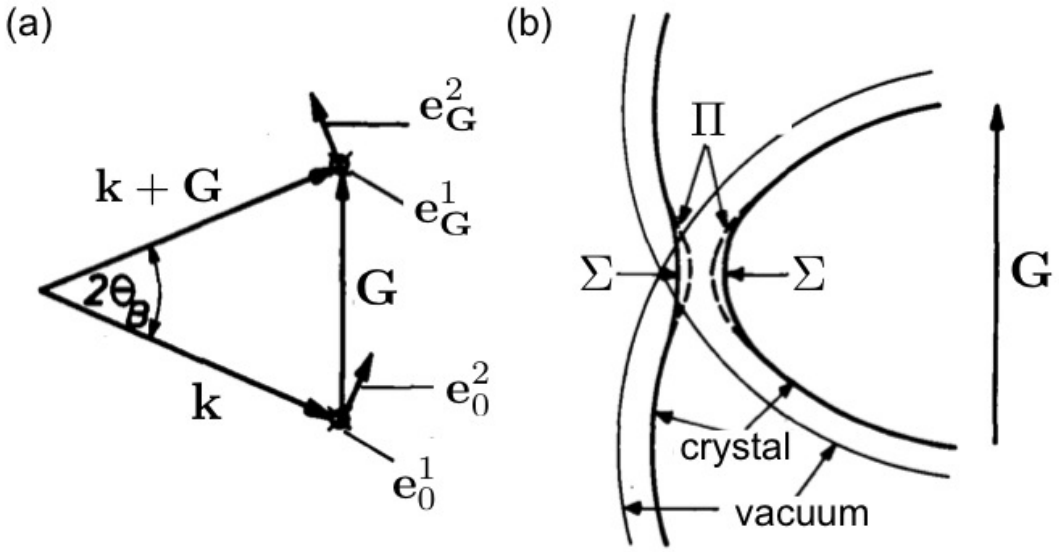


Fig. 11: *(a) Direction of polarization vectors. Vectors* $\mathbf{e}_0^1 = \mathbf{e}_{\mathbf{G}}^1$ *are normal to the plane of the plot. (b) Dispersion surface for* Σ *and* Π *polarizations. Taken from [11].*

Since $\nabla \cdot \mathbf{D} = 0$ the vectors $\mathbf{D_G}$ are orthogonal to $\mathbf{k}+\mathbf{G}$: $\mathbf{D_G} \cdot (\mathbf{k}+\mathbf{G}) = 0$. It can be shown that the corresponding component $\mathbf{H_G}$ is orthogonal both to $\mathbf{k}+\mathbf{G}$ and $\mathbf{D_G}$, while $\mathbf{E_G}$ lies in the plane of $\mathbf{D_G}$ and $\mathbf{k}+\mathbf{G}$, but the difference between $\mathbf{E_G}$ and $\mathbf{D_G}$ is very small owing to the smallness of χ.

For each Fourier component we can introduce two polarization vectors $\mathbf{e}_{\mathbf{G}}^{\sigma} \cdot (\mathbf{k}+\mathbf{G}) = 0$ and $\mathbf{e}_{\mathbf{G}}^{\sigma} \cdot \mathbf{e}_{\mathbf{G}}^{\sigma'} = \delta_{\sigma,\sigma'}$, where σ and σ' are either "1" or "2". Thus, we can write that $\mathbf{D_G} = \sum_{\sigma} D_{\mathbf{G}}^{\sigma} \mathbf{e}_{\mathbf{G}}^{\sigma}$. Since χ is very small, we can significantly simplify Eq. (56), arriving at the following system of equations for the scalar Fourier components of $\mathbf{D}$:

$$\left(k_0^2 - (\mathbf{k}+\mathbf{G})^2\right) D_{\mathbf{G}}^{\sigma} = \sum_{\mathbf{G}',\sigma'} \varkappa_{\mathbf{G}-\mathbf{G}'} \left(\mathbf{e}_{\mathbf{G}}^{\sigma} \cdot \mathbf{e}_{\mathbf{G}'}^{\sigma'}\right) D_{\mathbf{G}'}^{\sigma'}, \tag{58}$$

where

$$\varkappa_{\mathbf{G}-\mathbf{G}'} = -k_0^2 \chi_{\mathbf{G}-\mathbf{G}'} = \frac{4\pi r_e}{V_C} f_{\mathbf{G}-\mathbf{G}'} \tag{59}$$

with $f_{\mathbf{G}}$ as the atomic scattering factor, Eq. (15), and r_e as the classical electron radius of 2.82×10^{-13} cm. Note that since $\mathbf{e}_{\mathbf{G}}^{\sigma} \cdot \mathbf{e}_{\mathbf{G}'}^{\sigma'}$ is in general non-zero because $\mathbf{G} \neq \mathbf{G}'$, the latter equations couple both polarizations. Also note complete similarity between Eq. (58) and Eqs. (17) and (20) derived previously for the case of electron scattering in a crystal. Thus, $\varkappa_{\mathbf{G}-\mathbf{G}'}$ is analogous to the corresponding component of the crystal potential for electrons. The eigenenergies, or, bands $\omega_{\mathbf{k}}$ of X-ray Bloch states in a crystal are determined from the secular equation which is very much the same as Eq. (18). Analogously to the case of electrons the effect of absorption lies in complexity of the k-vectors, eigenenergies and Fourier components of the crystal "potential" $\varkappa$.

For scattering of X-rays from a finite crystal, the boundary conditions have to be formulated. Owing to the nature of Maxwell's equations given in the beginning of the section, and smallness of χ, the boundary conditions are very simple: both normal and tangential components of $\mathbf{D}$ have to be continuous across the boundary between the crystal and the vacuum. In general one

writes the fields in the vacuum before the slab (see Fig. 6) as:

$$\mathbf{D}(\mathbf{r}) = \hat{\mathbf{D}} e^{i\mathbf{k}_0\mathbf{r}} + \sum_{\mathtt{G}} \mathbf{R}_{\mathtt{G}} e^{i\mathbf{K}_{\mathtt{G}}^{-}\mathbf{r}}, \quad z < 0, \tag{60}$$

and after the slab:

$$\mathbf{D}(\mathbf{r}) = \sum_{\mathtt{G}} \mathbf{T}_{\mathtt{G}} e^{i\mathbf{K}_{\mathtt{G}}^{+}\mathbf{r}}, \quad z > d, \tag{61}$$

while in the crystal we seek the solution in terms of Bloch waves:

$$\mathbf{D}(\mathbf{r}) = \sum_{j} P_j \mathbf{D}_{\mathbf{k}_j}(\mathbf{r}), \tag{62}$$

where all the reflection ($\mathbf{R}_{\mathtt{G}}$), transmission ($\mathbf{T}_{\mathtt{G}}$) and P_j coefficients are found by satisfying the boundary conditions. It is noteworthy to remark, that, in analogy to electrons, the kinematic and dynamical theory of X-ray scattering can be also developed starting from the integral equation for the $\mathbf{D}$-field, analogous to Eq. (32):

$$\mathbf{D}(\mathbf{r}) = \hat{\mathbf{D}} e^{i\mathbf{k}_0\mathbf{r}} + \nabla \times \nabla \times \int d\mathbf{r}' \frac{e^{i\mathbf{k}_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \chi(\mathbf{r}') \mathbf{D}(\mathbf{r}'). \tag{63}$$

In analogy to electrons, if the scattering occurs away from Bragg condition, we have only one strong beam. If the Bragg reflection is excited, we have two strong beams $\mathbf{k}$ and $\mathbf{k}+\mathbf{G}$. In this case there is a natural choice for polarization vectors: (i) Σ-polarization for $\sigma = 1$: $\mathbf{e}_0^1 = \mathbf{e}_{\mathbf{G}}^1$ perpendicular to both $\mathbf{k}$ and $\mathbf{k}+\mathbf{G}$, and (ii) Π-polarization for $\sigma = 2$: $\mathbf{e}_0^2 = \mathbf{e}_{\mathbf{G}}^2$ in the plane of $\mathbf{k}$ and $\mathbf{k}+\mathbf{G}$, see Fig. 11. Via this particular choice of polarization vectors we can decouple the equations (58) for Σ and Π waves:

$$(k_0^2 - \varkappa_0 - k^2) D_0^\sigma = \varkappa_{\mathbf{G}} P_\sigma D_{\mathbf{G}}^\sigma, \tag{64}$$

$$(k_0^2 - \varkappa_0 - (\mathbf{k}+\mathbf{G})^2) D_{\mathbf{G}}^\sigma = \varkappa_{\mathbf{G}} P_\sigma D_0^\sigma, \tag{65}$$

where polarization factor $P_\sigma = \mathbf{e}_0^\sigma \cdot \mathbf{e}_{\mathbf{G}}^\sigma$ equals one for Σ-polarization, and $\cos 2\theta_B$ for Π-polarization. For each polarization the allowed k-vectors lie on the dispersion surface given by:

$$(k_0^2 - \varkappa_0 - k^2)(k_0^2 - \varkappa_0 - (\mathbf{k}+\mathbf{G})^2) = P_\sigma^2 |\varkappa_{\mathbf{G}}|^2. \tag{66}$$

As we can see, far away from the Bragg condition the dispersion surfaces are spheres around $\mathbf{k} = 0$ and $\mathbf{k} = \mathbf{G}$ with the radius $\sqrt{k_0^2 - \varkappa_0}$ for both polarizations. Near the Bragg condition the degeneracy between both polarizations is lifted and the smallest separation between the outer and inner branches becomes polarization-dependent:

$$\Delta k_\sigma = \frac{2\pi}{d_{ext}} = \frac{P_\sigma \varkappa_{\mathbf{G}}}{k_0 \cos\theta_B}, \tag{67}$$

see Figure 10. In this figure the dispersion surface in the vacuum, degenerate for both polarizations, is shown with thin lines, while the dispersion surface in the crystal is shown with thicker lines. The expressions for the coefficients D_0^σ and $D_{\mathbf{G}}^\sigma$ are exactly the same as those in Eq. (27), with $v_{\mathbf{G}}$ replaced by $P_\sigma \varkappa_{\mathbf{G}}$. The resulting Σ-fields are identical to those given by the Bloch waves ψ^I and ψ^{II} from the previous section. However, for Π-polarization we do not get pure sine and cosine waves but only a combination of both, since $\mathbf{e}_0^2 \neq \mathbf{e}_{\mathbf{G}}^2$. Correspondingly, only Σ-waves show an anomalous transmission effect, while Π-waves never vanish at the atomic positions completely. Interestingly, for multi-beam cases, important for X-ray diffraction, the situation complicates significantly, since both polarizations cannot be anymore decoupled.

6 Acknowledgments

I am grateful to Phivos Mavropoulos and Prof. Peter Dederichs.

References

[1] S. Blügel, *Scattering Theory: Born Series*, current book of manuscripts (A2)

[2] C. G. Darwin, Phil. Mag. **27**, 315; **27**, 675 (1914)

[3] P. P. Ewald, Ann. Physik **49**, 1; **49**, 117 (1916); **54**, 519 (1917)

[4] M. v. Laue, Ergeb. Exact. Naturw. **10**, 133 (1931)

[5] H. Bethe, Ann. Physik, **87**, 55 (1928)

[6] W. H. Zachariasen, *Theory of X-Ray Diffraction in Crystals* (Dover Publications, 1945)

[7] J. M. Cowley, *Diffraction Physics* (Elsevier, 1995)

[8] A. Authier, *Dynamical Theory of X-Ray Diffraction* (Oxford University Press, 2001)

[9] J. C. Slater, *Interaction of Waves in Crystals*, Reviews of Modern Physics **30**, 197 (1958)

[10] B. W. Batterman and H. Cole, *Dynamical Diffraction of X Rays by Perfect Crystals*, Reviews of Modern Physics **36**, 681 (1964)

[11] P. H. Dederichs, *Dynamical Diffraction Theory*, Berichte der Kernforschungsanlage Jülich, Nr. 797 (1971). Also published in *Solid State Physics* series, Vol. 27 (1972)

[12] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, (Brooks Cole, 1976)

[13] A. S. Davidov, *Quantum Mechanics*, (Pergamon Press, 1991)

A 4 Interaction of X-rays, Neutrons and Electrons with Matter

David P. DiVincenzo
PGI-2
Forschungszentrum Jülich GmbH

Contents

1 The Basics of Scattering Theory

Each of the scattering probes that we discuss here, be they particles or waves, permit, according to the tenets of quantum mechanics, a description of either sort. In fact, the wave theory is the best adapted as the unified framework that we will set up here. The incident beam will be treated as monoenergetic and unidirectional – and thus as a plane wave, with incident wave field

$$\Psi_{inc}(\mathbf{r}) = Ae^{i\mathbf{k}\cdot\mathbf{r}} \tag{1}$$

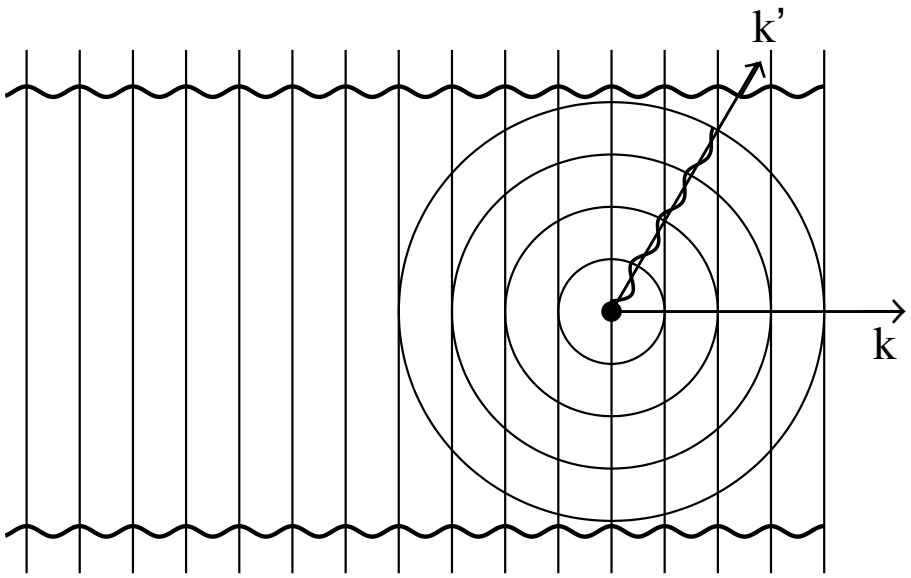


Fig. 1: A schematic of the scattering process from an atomic target. The incident plane wave (wavy line) has wavevector k; its constant-phase fronts are shown as straight lines. The scattered wave is an outgoing spherical wave (circles) going out in all directions, including in the wavevector direction k'.

The energy of the incident scatterer is a function of the magnitude of the wavevector $|\mathbf{k}|$; for nonrelativistic electrons or neutrons, $E = \hbar^2 k^2/2m$, and for light, $E = \hbar ck$. The direction of propagation is of course the direction of the vector $\mathbf{k}$, which will be conveniently described in spherical polar coordinates using angles $(\theta, \varphi) = \Omega$. We assume that there is a small scattering target fixed at the origin. In the relevant wave equation, this scatterer will be described by a potential energy function $V(\vec{r})$. The "interaction region" $|\mathbf{r}| < r_0$ is assumed to be the only region in which $V(\vec{r}) \neq 0$. Outside this interaction region the wave field also contains an outgoing spherical wave of the form

$$\Psi_{scat} = Af(\Omega)\frac{e^{ikr}}{r} \tag{2}$$

We have specialized to elastic scattering (appropriate for most of the scattering experiments considered in this chapter), so that the magnitude of the incident and scattered wavevectors k are the same.

The quantity $f(\Omega)$ is the central focus of our attention, describing the amount of scattering in the direction of the solid angle Ω. Note that the complex quantity f has units of length – it is a

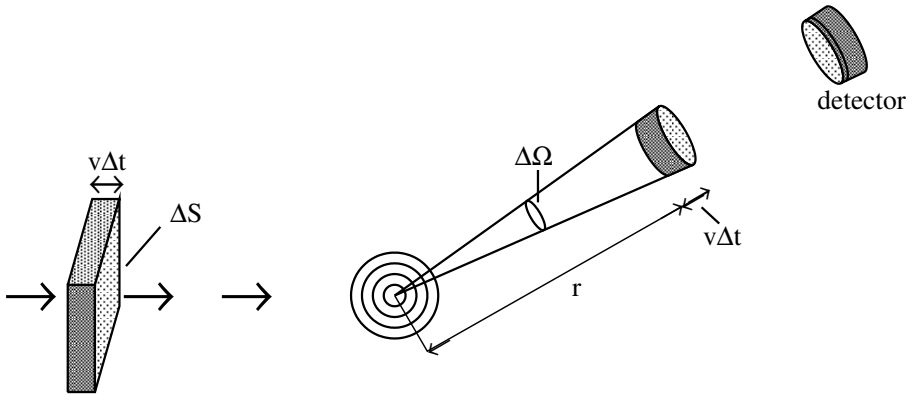


Fig. 2: Geometry of the scattering process.

"scattering length". It in fact directly indicates the normalized scattering flux $\Delta\sigma(\Omega)$ in a cone of solid angle $\Delta\Omega$ (see Fig. 2), in the direction Ω, for unit incident wave flux density:

$$\Delta\sigma = |f(\Omega)|^2 d\Omega \tag{3}$$

The total scattering cross section is

$$\sigma = \int |f(\Omega)|^2 d\Omega \tag{4}$$

We see that the phase of the complex scattering length $f(\Omega)$ does not appear in any of our expressions; however it is very important in the interference that occurs in scattering from two different scattering centers. This effect is beyond the scope of the present chapter.
For completeness, we note the other important quantity, the differential scattering cross section, which is simply the integrand of the quantity above:

$$\frac{d\sigma}{d\Omega} = |f(\Omega)|^2 \tag{5}$$

We end this section with a simple physical picture of the scattering cross section. Naturally, the above discussion implies that the full wave field is given by the sum of the incident and scattered waves, which is correct in the Born approximation:

$$\Psi_{tot}(\mathbf{r}) = A\left(e^{i\mathbf{k}\cdot\mathbf{r}} + f(\Omega)\frac{e^{ikr}}{r}\right) \tag{6}$$

This Born approximation expression does not take account of the fact that the flux of the incident beam is affected (and depleted) by scattering. The amount by which it is depleted is exactly the flux density through the area σ. One can have a simple picture of this result: the depleting effect of the scatterer is exactly the same as that of a fully absorbing screen with area σ (Figure 2). The common unit for σ in scattering physics is the *barn*, which, at $10^{-28}m^2$, is actually a large unit of area in many areas of particle and nuclear physics. The term originates from the 20th century American taunt to a poor thrower, "You couldn't hit the broad side of a barn."

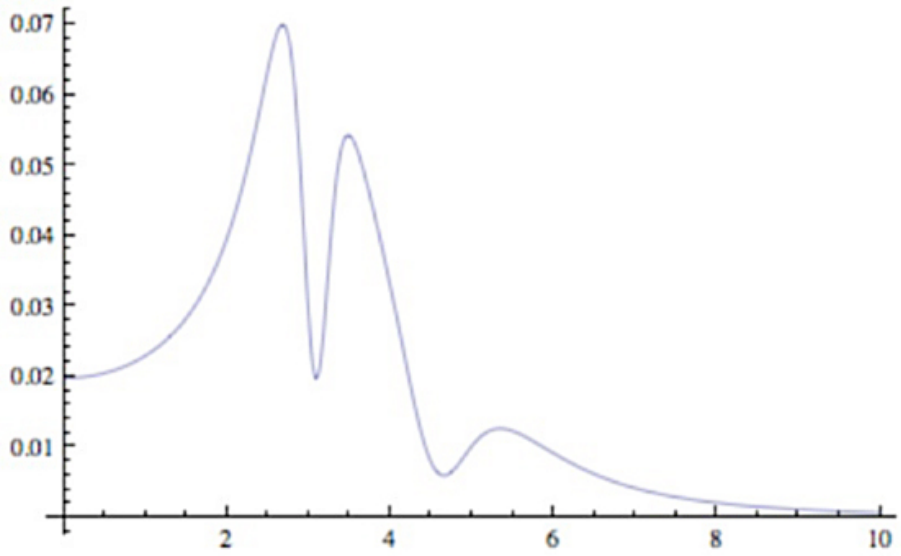


Fig. 3: A schematic of the square of ther real part of the polarizability $(\chi'(\omega))^2$ versus frequency ω. We see the low frequency, Rayleigh part (frequency independent), a complex intermediate frequency range in which anomalous dispersion and absorption occur, and then a high-frequency, Thompson part (going line $1/\omega^4$).

2 The Scattering of Electromagnetic Radiation from Atoms

While we will in this school primarily be concerned with high-energy scattering probes such as X-rays, we begin this discussion at the low-energy (that is, the low-frequency) end of the spectrum. An electromagnetic wave comprises transverse, perpendicular oscillating electric and magnetic fields. We first consider the effect of the electric fields on a target atom. At low frequencies, below that of any atomic resonances, the applied field will polarize the electrons bound to the atom, producing an electric polarization P proportional to the strength of the electric field:

$$P(\omega) = \chi(\omega)E(\omega) \tag{7}$$

At low frequencies, the electric polarizability of the atom $\chi(\omega)$ is independent of frequency ω. The resulting electric dipole oscillating with angular frequency ω, $P(\omega)e^{i\omega t}$, will radiate an outgoing spherical wave – this is the scattered wave of our general scattering theory. From classical electromagnetic theory, the efficiency with which this dipole radiates energy scales like the fourth power of the frequency; the net result for the scattering cross section is the formula for *Rayleigh scattering*:

$$\sigma_R(\omega) = \frac{8\pi}{3}\frac{\omega^4}{(4\pi\epsilon_0 c^2)^2}(\chi(0))^2 \tag{8}$$

Recall that this ω^4 dependence gives Rayleigh's explanation that the sky is blue.
Passing over the visible and ultraviolet region of the spectrum where atoms show complex resonant behavior in their scattering cross section (Figure 3), we consider a regime where the frequency is high enough that the binding of the electron to the atom is *irrelevant*; the electron oscillates as if it were in free space. In this regime a calculation of the oscillating dipole $P(\omega)$ is again straightforward, since it simply requires the calculation of the periodic displacement of a free particle subject to a sinusoidal force. The result in this regime is

$$\chi(\omega) = \frac{e^2}{m_e \omega^2} \tag{9}$$

The higher the frequency, the smaller the polarizability, because the electron has a shorter time in which to move. The scattering formula Eq. (8) still applies, so we get the simple, frequency-independent result for the cross section contribution per electron:

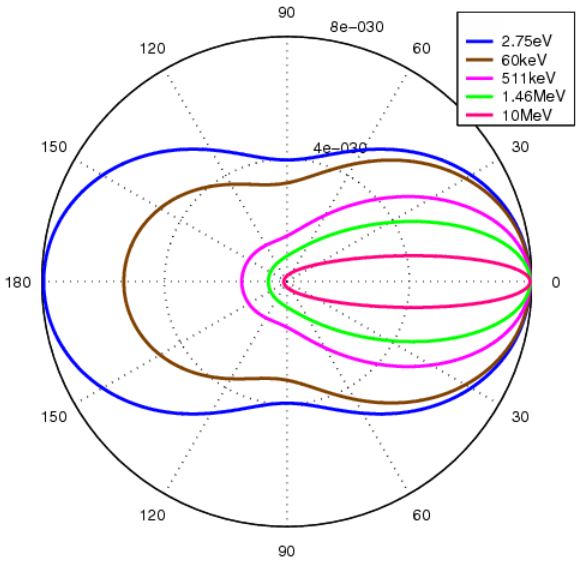


Fig. 4: Polar plot of the Klein-Nishina formula for the differential scattering cross section of X-rays by electrons. At low frequency the scattering goes to that of a classical dipole; at high frequency (the Compton regime) the cross section becomes more and more forward directed, best described as energy- and momentum-conserving photon-electron collisions. From [1].

$$\sigma_T = \frac{8\pi}{3} r_e^2 \tag{10}$$

This is the regime of *Thompson scattering*. Here

$$r_e = \frac{1}{4\pi\epsilon_0} \frac{e^2}{m_e c^2} \approx 2.8 \times 10^{-13} cm \tag{11}$$

is the so-called *classical electron radius*.

Even though the binding of the electron to the atomic nucleus is irrelevant in the Thompson scattering regime, it should be understood that, in the regime of low excitation intensity, the nevertheless remains associated with the atom, so long as the distance over which the electron travels under the influence of the time-oscillatory force is much smaller than the atomic radius. In this regime, X-ray scattering is non-destructive. Naturally, if the excitation intensity is raised to the point where this oscillation distance becomes comparable to or greater than the atomic

radius, we enter the regime of *high-intensity effects*, which can very realistically be achieved with strong X-ray sources such as the free-electron laser (FEL). In that case the X-ray probe is destructive, causing ionization and disruption of chemical structure, so the time available for this scattering probe to give useful information about condensed matter is limited.
Of course a more complete calculation is possible; the result for the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2}\left(1+\cos^2\theta\right). \tag{12}$$

The angular dependence appears in Fig. 3, showing the dipolar form that is also characteristic of the Rayleigh scattering regime. This figure shows the result of a much more general calculation due to Klein and Nishina [1], who calculated this scattering taking quantum and relativistic effects into account. The *Klein-Nishina formula* for the differential cross section is

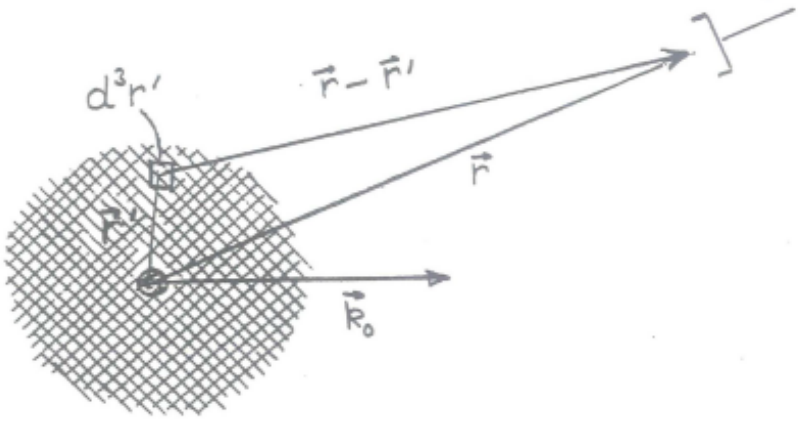


Fig. 5: Scattering geometry for discussion of atomic form factor.

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2}P(\omega,\theta)^2\left(P(\omega,\theta)+P(\omega,\theta)^{-1}-1+\cos^2\theta\right). \tag{13}$$

Here the factor

$$P(\omega,\theta) = \frac{1}{1+(\hbar\omega/m_ec^2)(1-\cos\theta)} \tag{14}$$

has a simple kinematical interpretation when we take the quantum point of view and consider the light to consist of particles (photons): it is the ratio of the photon energy after the scattering event to its original energy before scattering. Note that in the limit of small ω, $P(\omega,\theta) = 1$ and this expression reduces to the one for Thompson scattering. At high frequencies, when the photon energy $\hbar\omega$ becomes comparable to the rest energy of the electron $m_ec^2 = 511$keV, the scattering takes on a different character, and we enter the regime of *Compton scattering*. The scattering cross section becomes much more forward-directed, as we can see from the figure; the energetic photon suffers less and less of a deflection during the scattering from the electron, the higher its energy is.

3 Atomic Form Factor for X-rays

The Thompson scattering formula is clearly not the whole story of X-ray scattering from an atom. Even in the (considerable) frequency range in which The scenario for Thompson scattering applies (scattering from quasi-free electrons), we need to take account of the fact that the scattering is from the cloud of electrons that is bound to the atom. This means, in short, that the scatterers are not all at the origin of the coordinate system, and we must do a calculation to sum up their contributions.

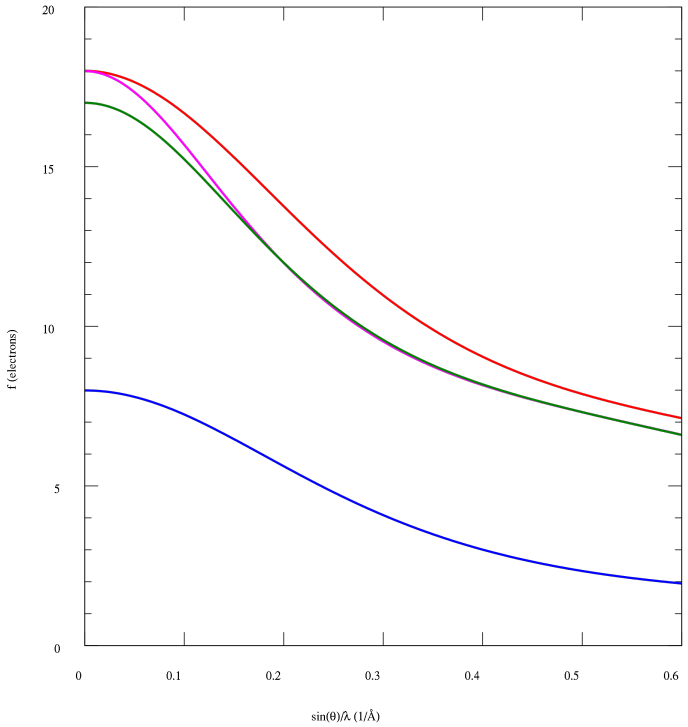


Fig. 6: The atomic form factor $f(Q)$ for several ions and elements, versus $\sin(\theta)/\lambda$. λ is wavelength, θ is scattering angle; with another 4π factor this expression is the scattering wavevector Q. Note that $f(0) = Z$ $Z =$ total number of electrons, not nuclear charge). From top to bottom these curves are for K^+, Cl^- (note that these have the same number of electrons), Cl and O. From [2].

Referring to the figure, we consider each volume element d^3r' to be a source of Thompson scattering with a strength governed by the probability that an electron is found in this volume element, which is given by the electron density function according to $n(\mathbf{r}')d^3r'$. The spherical wave that is emitted from that element involves the factor

$$n(\mathbf{r}')d^3r' \frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} e^{i\mathbf{k}_0\cdot\mathbf{r}'}. \tag{15}$$

Note that the final factor comes from the phase of the incident plane wave at the scattering point $\mathbf{r}'$. The overall scattering strength is given by integrating this quantity over the electron cloud:

$$\int d^3r' n(\mathbf{r}') \frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} e^{i\mathbf{k}_0\cdot\mathbf{r}'}. \tag{16}$$

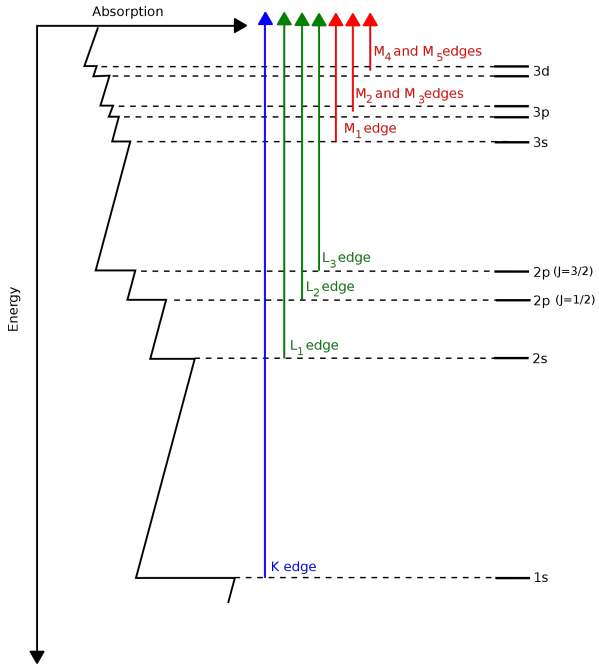


Fig. 7: Schematic of X-ray absorption structure, including many edges, for a typical atom with many inner shells. From [3].

Note that the angular factor from, e.g., the Thompson differential cross section formula will not appear inside this integral, since we will consider only the far field (i.e., $|\mathbf{r}| >> |\mathbf{r}'|$, so that the angular dependence can be put on as an overall factor once the integral is done.
Since $|\mathbf{r} - \mathbf{r}'| = r - \hat{\mathbf{r}} \cdot \mathbf{r}'$, we can approximate the integrand by

$$\frac{e^{ik_0|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} e^{i\mathbf{k}_0 \cdot \mathbf{r}'} \approx \frac{1}{r} e^{ik_0 r} e^{i(\mathbf{k}_0 - \mathbf{k}_1)\cdot \mathbf{r}'}. \tag{17}$$

Here $\mathbf{k}_1 = \hat{\mathbf{r}} k_0$. We see here appearing the *scattering wavevector*

$$\mathbf{Q} = \mathbf{k}_0 - \mathbf{k}_1 \tag{18}$$

With this we write our scattering amplitude

$$\int d^3r' n(\mathbf{r}') e^{i\mathbf{Q}\cdot\mathbf{r}'} \frac{e^{ik_0 r}}{r} = f_a(\mathbf{Q}) \frac{e^{ik_0 r}}{r}. \tag{19}$$

We identify the Fourier transform of the atomic electron density,

$$f_a(\mathbf{Q}) = \int d^3r' n(\mathbf{r}') e^{i\mathbf{Q}\cdot\mathbf{r}'}, \tag{20}$$

as the *atomic form factor* for X-ray scattering. It is a factor that must be accounted for in other applications of the scattering theory (e.g., for Bragg scattering). For example, it appears this way in the Thompson scattering differential cross section for an atom,

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2} |f_a(\mathbf{Q})|^2 \left(1 + \cos^2\theta\right) \tag{21}$$

We can see several features of the atomic form factor from Fig. 6. It is of course isotropic for atoms, so this it depends only on $|\mathbf{Q}|$. Its value at zero is very simple:

$$f_a(0) = Z \tag{22}$$

Z being the total electron number of the atom or ion. The figure shows two cases for which this number is the same, namely for the ions K^+ and Cl^-. The extension of these functions is the reciprocal of the extent of the atomic electron cloud in real space; thus we can observe that the Cl^- ion is considerably more extended than K^+.

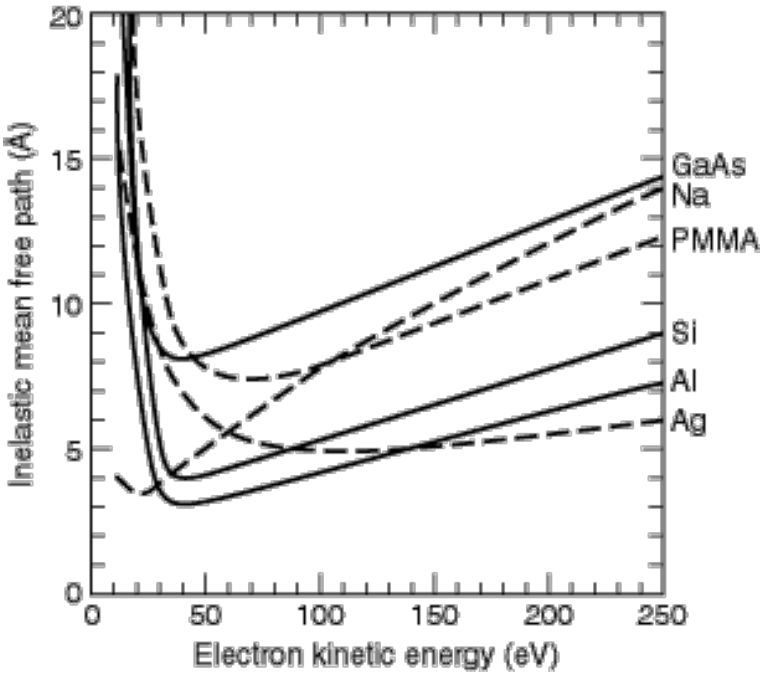


Fig. 8: Electron inelastic mean free path versus incident electron energy, for a range of materials. After [4].

4 X-ray Absorption and Dispersion

Our discussion above has so far ignored the phenomenon of absorption of radiation. We can trace this to our implicit assumption in Eq. (7) that the polarization vector is in-phase with the applied electric field, so that the polarizability function $\chi(\omega)$ is real. In fact the polarization has an out of phase component as well; elementary electromagnetic theory shows that a polarization oscillating out of phase with the electric field results in absorption of energy. Thus, we write the polarizability function as the sum of the real and an imaginary part:

$$\chi(\omega) = \chi'(\omega) + i\chi''(\omega) \tag{23}$$

For atoms, $\chi''(\omega)$ in the X-ray regime is fairly featureless, except for sharp *X-ray edges* that appear when the radiation can eject electrons from the inner electronic shells of the atom. Figure 7 shows the occurrence of these edges, and how they are interpreted in the shell model.

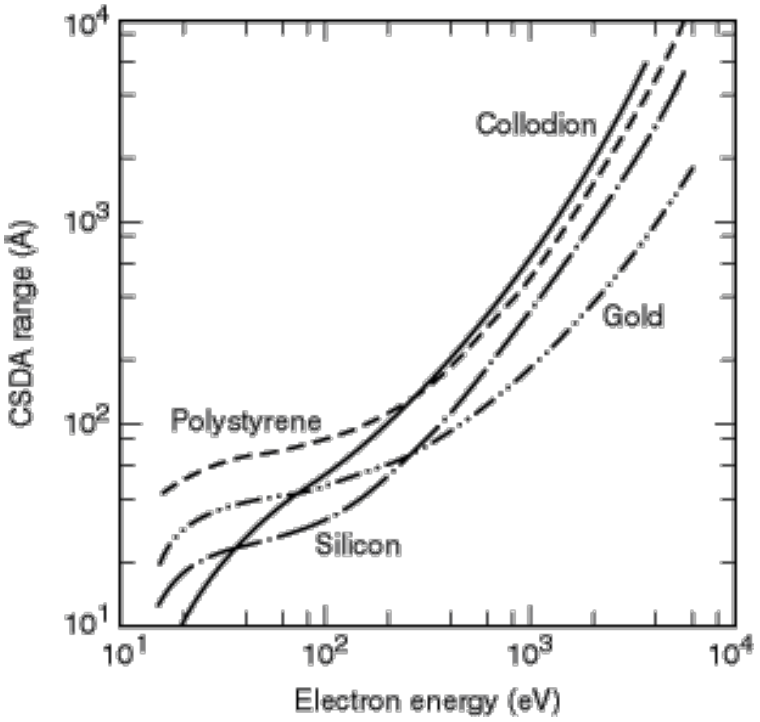


Fig. 9: Continuous Stopping Distance Approximation (CSDA) range of electrons vs. electron energy. After [4].

The occurrence of *dispersion*, significant variation in the lossless response $\chi'(\omega)$, is intimately tied to the appearance of structure in the lossy $\chi''(\omega)$ as exemplified by edges. This connection is embodied in the *Kramers-Kronig relations*. These relations for the $\chi'(\omega)$ function are

$$\chi'(\omega) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{\chi''(\omega)}{\omega' - \omega} d\omega', \tag{24}$$

$$\chi''(\omega) = -\frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{\chi'(\omega)}{\omega' - \omega} d\omega'. \tag{25}$$

The derivation of these relations is usually presented as an exercise involving Cauchy's theorem from complex analysis. I will take a moment to review a less rigorous but more physically informative demonstration involving only the elementary features of the Fourier transform. This derivation makes it more clear that the one and only one premise on which the Kramers-Kronig relations are based is the *causality* of the response of the system during scattering. Imagine that the scattering wave impinges as a wave packet on the scatterer, so that the time dependent electric field $E(t)$ becomes non-zero only after $t = 0$. The temporal polarization response is given using the Fourier transform of the polarizability function:

$$P(t) = \int_{-\infty}^{\infty} \chi(t - t') E(t') dt' \tag{26}$$

But because such a physical response is causal, $P(t) = 0$ if $t < 0$; the response cannot begin before the excitation has arrived. But for χ, this implies that

$$\chi(t) = 0 \text{ for } t < 0. \tag{27}$$

Now, we write $\chi(t) = \chi_e(t) + \chi_o(t)$, that is, we decompose χ into a sum of an even function of time and an odd function of time. Because of the causality condition these two functions are related:

$$\chi_e(t) = sgn(t)\chi_o(t) \tag{28}$$

Here the "sign function" $sgn(t)$ is $+1$ for $t \geq 0$ and -1 for $t < 0$. Fourier transforming this equation immediately gives the first Kramers-Kronig relation: The Fourier transform of χ_e is purely real and is in fact the real part of $\chi(\omega)$, $\chi'(\omega)$. The Fourier transform of the product is a convolution, the Fourier transform of $\chi_o(t)$ is $i\chi''(\omega)$, and the transform of $sgn(t)$ is $-i/(\pi\omega)$. The other relation is obtained similarly.

Hopefully this little discussion takes some of the mystery out of these relations. What do they have to do with X-ray scattering? We can see the connection by looking at the model of H. A. Lorentz for absorption based on a model of a damped resonator with resonant frequency ω_0. This is a good model for an electron bound in an atom; it captures only qualitatively the X-ray absorption edges, which involve not just the oscillation of the electron but also the ejection of the electrons into a continuum. But the "Lorentzian lineshape" for the absorption in Lorentz's model is very simple:

$$\chi''(\omega) = \frac{e^2}{m_e}\frac{\Gamma\omega}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \tag{29}$$

Here Γ is a linewidth or damping parameter. The Kramers-Kronig relation above requires that this absorbtion function be accompanied by the following frequency-dependent in-phase polarizability:

$$\chi'(\omega) = \frac{e^2}{m_e}\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2}. \tag{30}$$

Note that this expression intepolates between the two low-absorption regimes that we have discussed above: Rayleigh scattering (for $\omega \ll \omega_0$, and Thompson scattering for $\omega \gg \omega_0$. The Kramers-Kronig constraints say that there *must* be a regime of high loss in between, and that the in-phase polarizibility much also rise to a much higher value than in either of the two limits (in fact, $\chi'(\omega)_{max} = e^2/m(2\Gamma\omega_0 + \Gamma^2) \approx e^2/(2m\Gamma\omega_0)$ for $\Gamma \ll \omega_0$). This strong enhancement near an absorption feature of the real part of the polarizability, and therefore of the scattering cross section, and its strong frequency dependence, is known as *anomalous dispersion*. As you will learn elsewhere in this course, this phenomenon is used to enhance the contrast of one atomic element relative to another in X-ray scattering.

I conclude this section with a brief discussion of *magnetic X-ray scattering*. I have so far described the X-ray scattering process as involving only the electric field of the incident wave. Naturally, the electromagnetic wave also has a magnetic component, normal to the direction of propagation and also normal to the electric field. This magnetic field also induces a response, and causes an additional contribution to the scattered spherical wave. Most importantly, this scattering is sensitive to the magnetic state of the target – the scattering from an atom will be different when its spin is up or down. Thus, such contributions to the scattering can distinguish the magnetic state (ferromagnetic, antiferromagnetic, etc.) of a material. Generally, this magnetic contribution to the scattering is weak; the scattering amplitude has a prefactor $\hbar\omega/m_e c^2$, so that this scattering is generically suppressed for X-ray photon energies below 511keV. Wise use of magnetically-dependent anomalous dispersion can enhance the magnetic signal.

5 Electron Scattering

The basic physics of the scattering of electrons from matter is the same as that for X-ray photons: in the quantum theory the electrons have a wave description, and the basic scenario of scattering, in which there is an incident wave on the target, and an outgoing scattered spherical wave. The important qualitative distinction between electron scattering and X-ray scattering is that the strength of electron scattering is much greater than that of X-rays. Electrons will not penetrate a large thickness of material as X-rays will.

In fact, the scattering cross section for electrons can be deduced directly from the cross section for electromagnetic waves, already discussed above. In the electron wave equation (the Schroedinger equation), the scattering intensity from point $\mathbf{r}$ is determined by the potential function at that point $V(\mathbf{r})$. From a calculation of the scattering problem using this equation, the scattering form factor is given by the expression

$$f^e(Q) = \frac{2m_e e}{\hbar^2} \int_0^\infty \frac{V(r)\sin(qr)r^2 dr}{qr} \tag{31}$$

Note that by convention the form factor for electron scattering also contains the scattering length; this means that it has units of meters, rather than being dimensionless as the X-ray form factor is taken to be.

One further step permits f^e to be related directly to the X-ray form factor, since we can relate the scattering potential $V(\mathbf{r})$ to the electron density $n(r)$ whose fourier transform determines the $f(Q)$ for X-rays. This relation is via the Poisson equation, $\nabla^2 V(\mathbf{r}) = -\frac{e}{\epsilon_0} n(\mathbf{r})$. Fourier transforming this equation and substituting into Eq. (31) gives the *Mott-Bethe formula* for the electron form factor for an atom with atomic number Z:

$$f^e(Q,Z) = \frac{m_e e^2}{2\pi\hbar^2\epsilon_0}\left(\frac{Z - f(Q,Z)}{Q^2}\right) \tag{32}$$

This equation also includes the form factor Z/Q^2 for the atomic nucleus.

The principal item of practical interest that I will cover here is the theory of the stopping range of low-energy electrons in solid matter. For electrons with an incident energy in the range of 5 keV, the basic picture is that electrons slow down by a large sequence of scatterings in the material, each of which leads to a small loss of energy. We speak of the *continuous slowing down approximation* (CSDA) in calculating the electron range. This calculation again involves the polarizability of constituents $\chi(\omega)$. When summed over a large number of constituents, this response is called the *dielectric function* $\epsilon(Q,\omega)$; this expression singles out polarization leading to a scattering wavevector $\mathbf{Q}$. Then for an electron traveling with energy E, the probability of energy loss ω over a unit of distance is given by the expression

$$p(E,\omega) = \frac{m_e e^2}{\pi\hbar^2 E}\int_{q_-}^{q_+} \mathrm{Im}\left(\frac{-1}{\epsilon(Q,\omega)}\right)\frac{dQ}{Q} \tag{33}$$

Here $\hbar q_\pm = \sqrt{2mE} \pm \sqrt{2m(E-\hbar\omega)}$. $p(E,\omega)$ is known as the differential inverse mean free path. The stopping power $S(E)$, which is the energy loss per unit distance travelled along the electron path, is given by

$$S(E) = \int dE\hbar\omega p(E,\hbar\omega) \tag{34}$$

Finally, the distance over which the electron is stopped (actually, brought to a nominal kinetic energy of 10eV) is the CSDA range $R_0(E)$, given by

$$R_0(E) = \int_{10eV}^{E} \frac{dE'}{S(E')} \tag{35}$$

In Fig. 8 we show the inelastic mean free path for electrons in a wide variety of solid matierials, over the range of incident energies from 10 eV to 250 eV. This quantity continues to grow almost linearly above this energy, up to 2keV. There are two important things to note about this quantity: it has a minimum at a few tens of eV. Electrons are more penetrating at energies both above and below this. Second, the scale of this mean free path is very small, being a fraction of a nanometer over much of this energy range. Just a couple of atomic layers are effective at blocking the passage of a large fraction of electrons in this energy range.

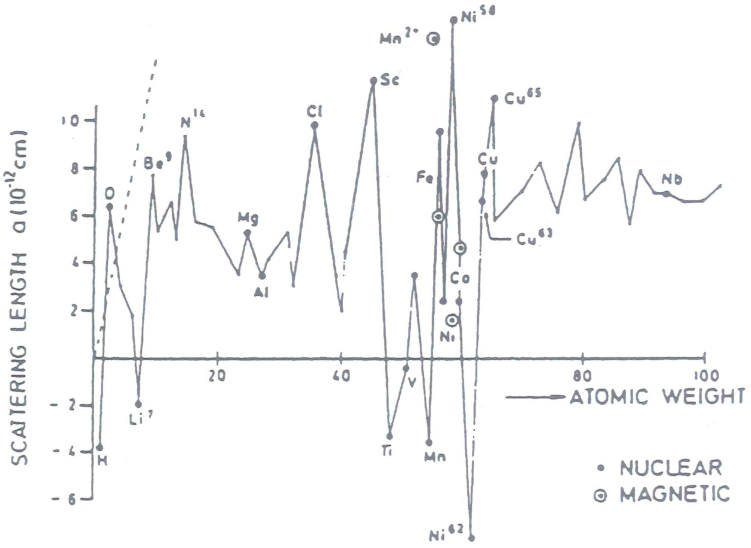


Fig. 10: Scattering length b for the atomic elements, showing the non-monotone dependence of b on atomic weight, even for isotopes of the same element. See [5].

Finally, Fig. 9 shows a sampling of the total travel range for incident electrons (note the much greater range of energies than in the previous figure) for several types of solid materials. Note that the total penetration range never exceeds 1 micrometer for any of the cases shown.

6 Neutron Scattering

The neutron, a particle with no charge, a mass very close to that of the proton, and a spin magnetic moment about 1000 times smaller than that of the electron, is a very useful scattering probe. Neutron beams of high intensity and sharply defined energy and direction can be produced and directed at targets; the neutron's lack of charge permits low-energy neutrons to penetrate deeply into matter. The free neutron is unstable; while its decay is fast compared with many radio-nuclei, at about 10 minutes this time is very long compared with that of the scattering process and detection, so that this decay can be ignored in discussions of neutron scattering.

The energy range of the scattering neutrons is usually "thermal", meaning that the kinetic energy of the neutrons is reduced by moderation (passage through a non-absorbing material) into the range of k_BT with $T \approx 300$K. (Moderation to lower energies is possible.) Monochromators pick off well defined energies from this moderated collection of neutrons.

Lacking an electric change, the neutron still has two means of interacting significantly with matter: first, its magnetic moment makes it sensitive to the magnetic scatterers in the target. Second, the strong nuclear force causes there to be a significant scattering cross section from each magnetic nucleus. It turns out that these contributions are roughly of the same order of magnitude, both are very important in the application of neutron scattering.

We will deal first with the scattering arising from the nuclear force. The strong interaction of the neutron with a many-nucleon atomic nucleus is very complex. However, the description we need of scattering is very much simpler, because at thermal energies, the wavelength of the quantum-mechanical (de Broglie) neutron wave is in the vicinity of 0.1 nm, comparable, in fact, to the internuclear spacings in molecules or solids. This wavelength is very long compared with the range of the strong nuclear force (about 10^{-6} nm). Thus, the neutron-nucleus interaction may be accurately represented as a delta-function at the origin; this is called the Fermi pseudopotential. Fermi writes

$$V(\mathbf{r}) = \frac{2\pi\hbar^2}{m_N} b\delta(\mathbf{r}) \tag{36}$$

Here m_N is the neutron mass. b has dimensions of a length, and is in fact the s-wave scattering length; it is also equal to the neutron form factor, since the delta-function form of the potential means that the form factor has no dependence on the scattering wavevector $\mathbf{Q}$.

In the simplest view b is just a simple scalar number. We must be a little bit more sophisticated, for several reasons. First, if the target nucleus possesses a non-zero nuclear spin quantum number I, then the scattering depends on the relative angles of the neutron spin vector $\mathbf{s}$ and the nucleus angular momentum vector $\mathbf{I}$; in general this spin dependence is quite strong. This effect is included by writing b as

$$b = b_c + \frac{2b_i}{\sqrt{I(I+1)}} \mathbf{s} \cdot \mathbf{I} \tag{37}$$

Thus, the scattering process takes two parameters to describe; for historical reasons, these parameters are called the *coherent* cross section b_c and the *incoherent* cross section b_i. In fact both parameters describe perfectly coherent wave scattering phenomena. However, it is typical in scattering experiments to have no control over the spin state of the target nucleus (there are now many exceptions to this); thus it has been traditional to consider the nuclear spin state to be random, causing the resulting scattering to be incoherent. We will state shortly the consequences of this for scattering from atomic crystals.

The second fact about b that we wish to note is that it also possesses an imaginary part ib". As with χ'' above, this constant describes the absorption of neutrons,due to nuclear reactions, during the scattering. Finally, the b "constants" can also be functions of energy. Generally b'' has a linear energy dependence, so that in tabulations the scattering energy must be stated. The real part is in most cases energy independent at thermal energies, although it should be noted that for a small minority of the atomic nuclei, there are already resonances, with anomalous dispersion and enhanced absorption, already at low energies.

All these parameters are accurately measured and can be found tabulated, typically with a couple of digits of accuracy but sometimes much more, for all the isotopes of the periodic table of elements. Unlike the X-ray and electron scattering lengths, which increase monotonically as

one moves down the periodic table, the b parameters, which depend on complex details of nuclear physics, are already large for the lightest nuclei, and vary tremendously from one element to the next, and vary to the same degree even for isotopes of the same element. We show this variation in Fig. 10. So, a crystal of pure He has a perfect periodic structure as seen either by X-ray scattering or neutron scattering (producing "Bragg peaks", see Chap B1), because pure helium consists almost entirely of one spinless isotope, He-4. But a crystal of pure selenium with equal amounts of Se-74 and Se-76 (these are both natural isotopes of Se, but these are not the natural abundances) looks highly disordered from the point of view of neutron scattering, producing a large component of non-Bragg diffuse scattering, because b_c for Se-74 and Se-76 are very different (0.8 and 12.2 barns, resp.). On the other hand, a pure crystal of arsenic looks disordered for the other reason; while there is only one stable isotope As-75, the four different permitted spin states of the I=3/2 As-75 nucleus scatter with considerably different strengths (because $b_i = -0.7$ barns).

7 Magnetic Neutron Scattering

The neutron is a chargeless particle, but it has a magnetic moment, which is about the same in magnitude as the protons, and about 100 times smaller than that of the electron. We analyse the scattering of the neutron from the field arising from the spin of an electron at position $\mathbf{r}'$:

$$\mathbf{B}_S = \frac{\mu_0}{4\pi}\nabla_{\mathbf{r}} \times \left(\boldsymbol{\mu}_e \times \nabla_{\mathbf{r}} \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right), \tag{38}$$

here $\boldsymbol{\mu}_e = g_e \mu_B \mathbf{s}_e$ is the electron spin magnetic moment; the Bohr magneton is μ_B and the electron spin operator is $\mathbf{s}_e$. The field arising from electron orbital motion is

$$\mathbf{B}_L = -\frac{\mu_0 e}{4\pi} \frac{\mathbf{v}_e \times (\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} \tag{39}$$

(we use the Biot-Savart law for a particle with charge $-e$ and velocity $\mathbf{v}_e$). Here we further consider only the spin field, yielding the potential

$$V(\mathbf{r}) = -\boldsymbol{\mu}_n \mathbf{B}_S = -\boldsymbol{\mu}_n \frac{\mu_0 g_e \mu_B}{4\pi} \nabla_{\mathbf{r}} \times \int \mathrm{d}\mathbf{r}' \mathbf{s}_e(\mathbf{r}') \times \nabla_{\mathbf{r}} \frac{1}{|\mathbf{r}-\mathbf{r}'|}, \tag{40}$$

$\boldsymbol{\mu}_n$ is the neutron magnetic moment, and $\mathbf{s}_e(\mathbf{r}')$ is the electronic spin density. The scattering amplitude requires a calculation of a double integral

$$I = \int \mathrm{d}\mathbf{r} \mathrm{e}^{-\mathrm{i}\mathbf{Q}\mathbf{r}} \nabla_{\mathbf{r}} \times \int \mathrm{d}\mathbf{r}' \mathbf{s}_e(\mathbf{r}') \times \nabla_{\mathbf{r}} \frac{1}{|\mathbf{r}-\mathbf{r}'|} \tag{41}$$

We perform this evaluation in the Furier domain

$$\frac{1}{r} = \frac{1}{2\pi^2} \int \mathrm{d}\mathbf{q} \frac{\mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{r}}}{q^2}.$$

One obtains

$$\begin{aligned} I &= -\frac{1}{2\pi^2} \int \mathrm{d}\mathbf{r}' \int \mathrm{d}\mathbf{r} \mathrm{e}^{-\mathrm{i}\mathbf{Q}\mathbf{r}} \int \mathrm{d}\mathbf{q}\hat{\mathbf{q}} \times \mathbf{s}_e(\mathbf{r}') \times \hat{\mathbf{q}} \mathrm{e}^{\mathrm{i}\mathbf{q}(\mathbf{r}-\mathbf{r}')} \\ &= -4\pi \hat{\mathbf{Q}} \times \int \mathrm{d}\mathbf{r}' \mathbf{s}_e(\mathbf{r}') \times \hat{\mathbf{Q}} \mathrm{e}^{-\mathrm{i}\mathbf{Q}\mathbf{r}'}. \end{aligned} \tag{42}$$

Note that the application of $\nabla_{\mathbf{r}}$s to the exponential term $\mathrm{e}^{\mathrm{i}\mathbf{q}(\mathbf{r}-\mathbf{r}')}$ results in the simple factors $\mathrm{i}\mathbf{q} = \mathrm{i}q\hat{\mathbf{q}}$. The integration over $\mathbf{r}$ gives the delta function $\delta(\mathbf{Q}-\mathbf{q})$, permitting the integration over $\mathbf{q}$ to be completed. Specializing to the case of constant spin direction $\mathbf{s}_e(\mathbf{r}') = s_e(\mathbf{r}')\hat{\mathbf{s}}$ we get a scattering amplitude

$$f(\mathbf{Q}) = -\mu_n \frac{2m_n}{\hbar^2}\frac{\mu_0 g_e \mu_B}{4\pi}\hat{\mathbf{Q}}\times\hat{\mathbf{s}}\times\hat{\mathbf{Q}}F_{magn}(\mathbf{Q}), \tag{43}$$

here we see the *magnetic form factor*

$$F^{magn}(\mathbf{Q}) = \int \mathrm{d}\mathbf{r}'\mathrm{e}^{\mathrm{i}\mathbf{Q}\mathbf{r}'}s_e(\mathbf{r}'). \tag{44}$$

Using the usual cross-product identity $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$ one gets $\hat{\mathbf{Q}}\times\hat{\mathbf{s}}\times\hat{\mathbf{Q}} = \hat{\mathbf{s}} - (\hat{\mathbf{s}}\cdot\hat{\mathbf{Q}})\hat{\mathbf{Q}}$; note that this is the component of $\hat{\mathbf{s}}$ perpendicular to $\hat{\mathbf{Q}}$. Thus the scattering amplitude (44) is related to the Fourier transform of the spin density component perpendicular to the scattering vector $\mathbf{Q}$. So, magnetic neutron scattering allows a determination of both the size and the direction of the magnetisation in a material of interest. Originally neutron scattering was the only practical probe for the determination of the magnetic structure of solids. In the present time magnetic X-ray scattering with X-rays produced using synchrotron radiation sources can also deliver such information.

Let us estimate the magnitude of the magnetic scattering length (43). The neutron magnetic moment is $\boldsymbol{\mu}_n = \frac{1}{2}g_n\mu_N\boldsymbol{\sigma}$, where g_n is the neutron g-factor, μ_N the nuclear magneton and $\boldsymbol{\sigma} = 2\mathbf{s}_n$ is the Pauli spin operator. We can estimate the prefactor in (43) to be

$$\frac{2m_n}{\hbar^2}\frac{\mu_0}{4\pi}g_n\frac{e\hbar}{2m_p}g_e\frac{e\hbar}{2m_e} \approx 4\frac{e^4}{(4\pi\epsilon_0\hbar c)^2}\frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 4\alpha^2 a_0,$$

where $g_n \approx -4$, $g_e \approx -2$, $m_n \approx m_p$, $\mu_0 = 1/\epsilon_0 c^2$, the Bohr magneton $\mu_B = e\hbar/2m_e$, the nuclear magneton $\mu_N = e\hbar/2m_p$, the fine structure constant $\alpha = e^2/4\pi\epsilon_0\hbar c$ and the Bohr radius $a_0 = 4\pi\epsilon_0\hbar^2/m_e e^2$ have been used. Note that $\alpha^2 a_0$ is the classical electron radius r_e; this happens to be in the same range as nuclear scattering lengths b. Thus the nuclear and magnetic scattering are of competitive size (very much unlike the X-ray case); this means that interference between the two forms of scattering can in practice occur.

References

[1] Wikipedia article, “Klein-Nishina Formula”.

[2] Wikipedia article, “Atomic Form Factor”.

[3] Wikipedia article, “X-ray Absorption Spectroscopy”.

[4] Piero Pianetta, *X-ray data booklet*, http://xdb.lbl.gov/Section3/Sec_3-2.pdf.

[5] Varley Sears, *Neutron Cross Sections*, Neutron News **3** (3), 26 (1992).

A 5 Correlation Functions Measured by Scattering Experiments

R. Zorn

Jülich Centre for Neutron Science

Forschungszentrum Jülich GmbH

Contents

[0]Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

The central statement of this lecture is that scattering experiments indirectly measure correlation functions. The usual derivation of scattering laws is based on the fact that the scattering law (for neutrons, photons or any other radiation) is essentially the absolute square of the Fourier transform of a scattering density. In Fig. 1 this is shown as the left way from the density $\rho(\mathbf{r})$ to $S(Q)$. The Wiener-Khintchine theorem

$$|\mathcal{F}[f(x)]|^2 = \mathcal{F}[\langle f(0)f(x)\rangle] \tag{1}$$

(with the Fourier transform $\mathcal{F}$ defined as in the appendix) now states that the absolute square of a Fourier transform is the Fourier transform of the autocorrelation function. This opens another way (the right one in Fig. 1) to calculate the scattering law. Apart from elucidating the meaning of the scattering law in another way, this gives an alternative to calculate it even if the density itself is not known.

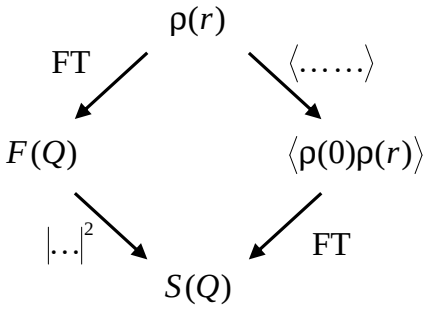


Fig. 1: *The two ways to calculate the scattering law from the microscopic density, left: as the absolute-squared Fourier transform of the density, right: as the Fourier transform of the correlation function.*

This lecture will in the first section treat the results from a static system where the scattering is completely elastic. In this situation the scattering will only contain information about the structure. Strictly speaking, this is a fictitious assumption because all materials show some dynamics (quantum-mechanically even at zero temperature). Nevertheless, the broad range of diffraction methods is covered with sufficient accuracy. The second part of the lecture will deal with inelastic scattering. In this experiment, scattering gives information about the structure via the momentum transfer and about the dynamics via the energy transfer. For inelastic scattering a Fourier transform in time has to be carried out in addition leading from the time correlation function to the scattering function.
Textbooks on scattering theory usually restrict themselves to a single probe. For neutron scattering refs. 1–3 can be recommended, for light scattering refs. 4–8, and for x-ray scattering refs. 9–11.

2 Scattering from static systems

In this section it will be assumed that the scattering system is *static*. It is either represented by fixed positions of point scatterers in space, $\mathbf{r}_j$, or a time-independent density, $\rho(\mathbf{r})$. The former case can be included in the latter by considering the microscopic density as a sum of

delta functions:

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j)\,. \tag{2}$$

From the fact that the scatterers are fixed follows that the scattering will be *elastic*, i.e. the energy of the scattered particles will not change due to the scattering process. This is clear from classical mechanics because a system which is static before and after the scattering process cannot exchange energy. The equivalent argument from the wave picture would be that upon scattering by fixed centres there is no Doppler shift of the frequency.

2.1 Structure factor from density

The result of an elastic scattering experiment is usually expressed in terms of the *differential cross-section* which is the probability density that a particle is scattered into a solid angle element $\mathrm{d}\Omega$ normalised to the intensity of the incident beam:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left\langle \left| \sum_{j=1}^{N} b_j \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_j) \right|^2 \right\rangle . \tag{3}$$

b_j is a measure of the 'scattering power' of the particle. From the dimensions it is obvious that it has the dimension [length]. Therefore, b_j is called the *scattering length*. Note that the scattering length is not necessarily positive. $b_j < 0$ just means that scattering leads to a reversal of the amplitude, in other words a phase shift π. The scattering length may even be complex. In that case, the imaginary part corresponds to absorption of the scattered particle by the scatterer.
It can be seen that expression (3) does not contain the scattering angle 2θ directly but a *scattering vector* $\mathbf{Q}$. It is the vectorial difference of the wave vector $\mathbf{k}'$ after scattering and that before scattering, $\mathbf{k}$. The wave vectors are defined by having the length $|\mathbf{k}| = k = 2\pi/\lambda$ and the direction of the propagation of the wave. For elastic scattering $k' = k$, and the definition of $\mathbf{Q}$ is graphically demonstrated by the black (isosceles) triangle in Fig. 2 resulting in

$$Q = \frac{4\pi}{\lambda}\sin\theta\,. \tag{4}$$

From this equation one can see that scattering depends on a combination of the scattering angle and the wavelength of the scattered radiation. The same Q can be obtained by different combinations of 2θ and λ. Quantum-mechanically Q corresponds to the momentum transfer due to the scattering process: $\hbar\mathbf{Q} = \hbar\mathbf{k}' - \hbar\mathbf{k} = \mathbf{p}' - \mathbf{p}$.
At this point it is necessary to explain the meaning of the average $\langle \ldots \rangle$ in (3) and justify it. Of course for a completely arrested system and completely coherent radiation, (3) would be valid without the average. Experimentally, this situation is only realised in laser light scattering from rigid objects. There, the experiments as well as the calculation do not yield a smooth function $\mathrm{d}\sigma/\mathrm{d}\Omega$ but an assembly of so-called speckles. For two reasons this situation is exceptional and the observed scattering is usually an average:

1. If a dynamics exists, even if it is sufficiently slow not to cause a noticeable inelasticity, the particles will rearrange over the duration of the experiment. In this sense, $\langle \ldots \rangle$ expresses a temporal average over the experimental time.

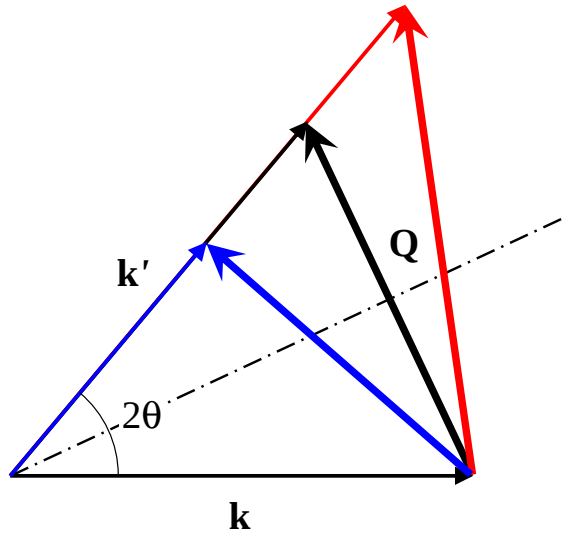


Fig. 2: *Definition of the scattering vector* $\mathbf{Q}$ *in terms of the incident and final wave vectors* $\mathbf{k}$ *and* $\mathbf{k}'$*. The black (isosceles) triangle corresponds to elastic scattering. The blue and red ones correspond to inelastic scattering with energy loss or gain of the scattered radiation, respectively.*

2. If the radiation used is not highly coherent, the sum over the amplitudes in (3) has to be restricted to the coherence volume which is usually much smaller than the sample volume. The results from the individual regions have to be added as intensities, i.e. after the absolute-square. This implies the same average but to be interpreted as a thermodynamic average over different realisations of the particle positions. In the case of ergodic systems both averages have the same result.

With the assumption that all scatterers are identical (for neutron scattering implying that they are the same isotope and have the same spin orientation) one can factor out the material-specific properties N and b:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |b|^2 N S(\mathbf{Q})\,. \tag{5}$$

with the remaining term

$$S(\mathbf{Q}) = \frac{1}{N}\left\langle \left| \sum_{j=1}^{N} \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_j) \right|^2 \right\rangle \tag{6}$$

which depends solely on the statistics of the positions of the scatters. $S(\mathbf{Q})$ is called *structure factor*.
If the scattering is not effected by individual scatterers but by a field (e.g. the magnetic field for neutrons) or a distribution (e.g. the electron density for x-rays) one has to use a continuum description instead of (6):

$$S(\mathbf{Q}) = \frac{1}{N}\left\langle \left| \int_V \mathrm{d}^3 r \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})\rho(\mathbf{r}) \right|^2 \right\rangle\,. \tag{7}$$

It is easy to verify that this expression corresponds to (6) with the definition (2) of the microscopic density inserted. Expression (7) is the absolute square of the Fourier transform of the density and thus represents the ‘left way’ in Fig. 1.

But even if the individual scatterers are point-like, the continuum description may be useful if their exact positions are not known but only their mesoscopic densities. In this case expression (7) will be a good approximation as long as the length scale defined by Q is large compared to the distances between the scatterers, $Q \ll 2\pi/\text{distance}$ (e.g. for small-angle x-ray or -neutron scattering).
At that point a simple way to introduce mixed scatterers is to start with the *scattering length density*

$$\rho_b(\mathbf{r}) = \sum_{j=1}^{N} b_j \delta(\mathbf{r} - \mathbf{r}_j) \,. \tag{8}$$

instead of the density. By including the scattering properties in the density, equation (5) can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left\langle \left| \int_V \mathrm{d}^3 r \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}) \rho_b(\mathbf{r}) \right|^2 \right\rangle \,. \tag{9}$$

Thus, the differential cross section is the absolute square of the Fourier transform of the scattering length density. For neutron scattering, this concept is used to obtain a low-resolution description for small-angle scattering and reflectometry. For light scattering the local dielectric constant of the medium plays the rôle of $\rho_b(\mathbf{r})$.

2.2 Structure factor from pair correlation function

The second way to derive the scattering law starts with applying the definition of the absolute square, $|X|^2 = X^* X$ to equation (6):

$$\begin{aligned} S(\mathbf{Q}) &= \frac{1}{N} \left\langle \left(\sum_{j=1}^{N} \exp(-\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_j) \right) \left(\sum_{k=1}^{N} \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_k) \right) \right\rangle \\ &= \frac{1}{N} \sum_{j,k=1}^{N} \langle \exp(\mathrm{i}\mathbf{Q}\cdot(\mathbf{r}_k - \mathbf{r}_j)) \rangle \,. \end{aligned} \tag{10}$$

From this expression two characteristic properties of scattering become clear:

1. The scattering law arises from particle pairs (j, k).

2. Only distances between particles enter the expression, not the individual positions. The scattering law remains invariant under translation of the whole sample.

In order to proceed in a similar way as before, we introduce the two-particle density

$$\rho(\mathbf{r}_1)\rho(\mathbf{r}_2) = \sum_{j,k=1}^{N} \delta(\mathbf{r}_1 - \mathbf{r}_j)\delta(\mathbf{r}_2 - \mathbf{r}_k) \tag{11}$$

which is the joint probability that particle j is found at $\mathbf{r}_1$ and particle k at $\mathbf{r}_2$. It is important that in general the average of this probability density is not just the product of the average densities:

$$\langle \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) \rangle \neq \langle \rho(\mathbf{r}_1) \rangle \langle \rho(\mathbf{r}_2) \rangle = {\rho_0}^2 \tag{12}$$

($\rho_0 = N/V$). The reason for this is that usually there is an interaction between particles which enhances or reduces the probability for particles close to each other. E.g. if one imagines particles with a hard core of radius R then $\langle\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\rangle$ vanishes for all $\mathbf{r}_1$ and $\mathbf{r}_2$ which would imply a 'collision' of the particles, $0 < |\mathbf{r}_2 - \mathbf{r}_1| < 2R$. Nevertheless, in a translationally invariant system one of the positions can be chosen arbitrarily, especially as the origin, so that

$$\langle\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\rangle = \langle\rho(\mathbf{0})\rho(\mathbf{r}_2 - \mathbf{r}_1)\rangle = \rho_0 \left\langle \sum_{j,k=1}^{N} \delta(\mathbf{r}_j - \mathbf{r}_k + \mathbf{r}_2 - \mathbf{r}_1) \right\rangle . \tag{13}$$

For a system of identical scatterers the two-particle density (11) can now be used to express the structure factor:

$$\begin{aligned} S(\mathbf{Q}) &= \frac{1}{N} \left\langle \int_V \mathrm{d}^3 r_1 \int_V \mathrm{d}^3 r_2 \exp(\mathrm{i}\mathbf{Q}\cdot(\mathbf{r}_2 - \mathbf{r}_1))\rho(\mathbf{r}_1)\rho(\mathbf{r}_2) \right\rangle \\ &= \frac{1}{\rho_0} \int_{V_d} \mathrm{d}^3 r \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle . \end{aligned} \tag{14}$$

Note that in this last expression $\mathbf{r}$ does not have the meaning of an absolute position but that of a vectorial distance and consequently the volume of integration V_d is not the sample volume V but the volume of possible distances within the sample.
In the literature often alternative definitions of the pair correlation function are used (instead of using $\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle$ directly). The most common definition in the context of liquids and colloids is

$$g(\mathbf{r}) = \frac{\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle}{{\rho_0}^2} - \frac{\delta(\mathbf{r})}{\rho_0} . \tag{15}$$

The normalisation by ${\rho_0}^2$ has the effect that for non-interacting particles or at distances where the interaction is weak, $g(\mathbf{r}) = 1$. The subtraction of the delta function removes the singularity of $\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle$ at $\mathbf{r} = \mathbf{0}$ due to the $j = k$ terms in (11). With this pair correlation function the structure factor can be written as

$$S(\mathbf{Q}) = 1 + \rho_0 \int_{V_d} \mathrm{d}^3 r \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})(g(\mathbf{r}) - 1) . \tag{16}$$

Here, the $1+$ compensates the delta function term subtracted in (15). In addition, one usually writes $g(\mathbf{r}) - 1$ instead of simply $g(\mathbf{r})$ in the Fourier transform. This avoids a delta function term arising in the limit $V_d \to \infty$ at $Q = 0$. In that limit, this 'trick' only changes the result at $Q = 0$ which is the (unobservable) forward scattering. Nevertheless, strictly speaking, one loses the scattering contribution by the overall sample shape. But this only affects the very low Q region if the sample has macroscopic dimensions $\gg 2\pi/Q$.
In many physical systems the interaction between particles is not directional with the consequence that $g(\mathbf{r})$ depends only on the distance $r = |\mathbf{r}|$. In this case by symmetry follows that also $S(\mathbf{Q})$ is only a function of $Q = |\mathbf{Q}|$. Formula (131) from the appendix can be applied resulting in

$$S(Q) = 1 + \frac{4\pi\rho_0}{Q} \int_0^\infty (g(r) - 1)\sin(Qr) r \mathrm{d}r . \tag{17}$$

Another definition, more often used for crystalline structures, is the Patterson function:

$$P(\mathbf{r}) = \frac{\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle}{\rho_0} \tag{18}$$

which Fourier-transforms into the structure factor without further prefactors:

$$S(\mathbf{Q}) = \int_{V_d} \mathrm{d}^3 r \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}) P(\mathbf{r})\,. \tag{19}$$

2.3 Example: crystalline lattice

We first consider the case of an orthorhombic lattice with a single point scatterer at the origin of the unit cell. With the lattice parameters a_1, a_2, and a_3 the positions of the atoms are given as

$$\mathbf{r}_{uvw} = (ua_1, va_2, wa_3)\,. \tag{20}$$

$u = 1\ldots U-1$, $v = 1\ldots V-1$, and $w = 1\ldots W-1$ are the indices of the lattice positions. The dimensions of the crystal are $Ua_1 \times Va_2 \times Wa_3$ and the total number of scatterers is $N = UVW$.
With these informations the complex amplitude in the expression for the structure factor (6) can readily be calculated:

$$\begin{aligned}
F(\mathbf{Q}) &\equiv \sum_{j=1}^{N} \exp\left(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_j\right) \\
&= \sum_{u=0}^{U}\sum_{v=0}^{V}\sum_{w=0}^{W} \exp\left(\mathrm{i}\mathbf{Q}\cdot(uQ_xa_1 + vQ_ya_2 + wQ_za_3)\right) \\
&= \sum_{u=0}^{U} \exp\left(\mathrm{i}uQ_xa_1\right) \sum_{v=0}^{V} \exp\left(\mathrm{i}vQ_ya_2\right) \sum_{w=0}^{W} \exp\left(\mathrm{i}wQ_za_3\right) \\
&= \sum_{u=0}^{U} \left(\exp\left(\mathrm{i}Q_xa_1\right)\right)^u \sum_{v=0}^{V} \left(\exp\left(\mathrm{i}Q_ya_2\right)\right)^v \sum_{w=0}^{W} \left(\exp\left(\mathrm{i}Q_za_3\right)\right)^w .
\end{aligned}$$

The last step identifies the three sums as geometric series which can be calculated in closed form:

$$\begin{aligned}
F(\mathbf{Q}) &= \frac{1-\exp\left(\mathrm{i}UQ_xa_1\right)}{1-\exp\left(\mathrm{i}Q_xa_1\right)} \cdots \\
&= \frac{\exp\left(-\mathrm{i}UQ_xa_1/2\right) - \exp\left(\mathrm{i}UQ_xa_1/2\right)}{\exp\left(-\mathrm{i}Q_xa_1/2\right) - \exp\left(\mathrm{i}Q_xa_1/2\right)} \frac{\exp\left(\mathrm{i}UQ_xa_1/2\right)}{\exp\left(\mathrm{i}Q_xa_1/2\right)} \cdots \\
&= \frac{\sin\left(UQ_xa_1/2\right)}{\sin\left(Q_xa_1/2\right)} \exp\left(\mathrm{i}(U-1)Q_xa_1/2\right) \ldots .
\end{aligned}$$

Because $|\exp(\mathrm{i}x)| = 1$ for real x, the complex exponential drops out upon calculating the structure factor:

$$S(\mathbf{Q}) = \frac{1}{N}|F(\mathbf{Q})|^2 = \frac{1}{N} \frac{\sin^2\left(UQ_xa_1/2\right)}{\sin^2\left(Q_xa_1/2\right)} \frac{\sin^2\left(VQ_ya_2/2\right)}{\sin^2\left(Q_ya_2/2\right)} \frac{\sin^2\left(WQ_za_3/2\right)}{\sin^2\left(Q_za_3/2\right)}\,. \tag{21}$$

So the structure factor of the orthorhombic crystal consists of three identical factors, each one depending on a single Cartesian component of the scattering vector $\mathbf{Q}$. The functional form is called Laue function, shown in Fig. 3 exemplarily for $U = 5$ and $U = 10$. It can be easily seen

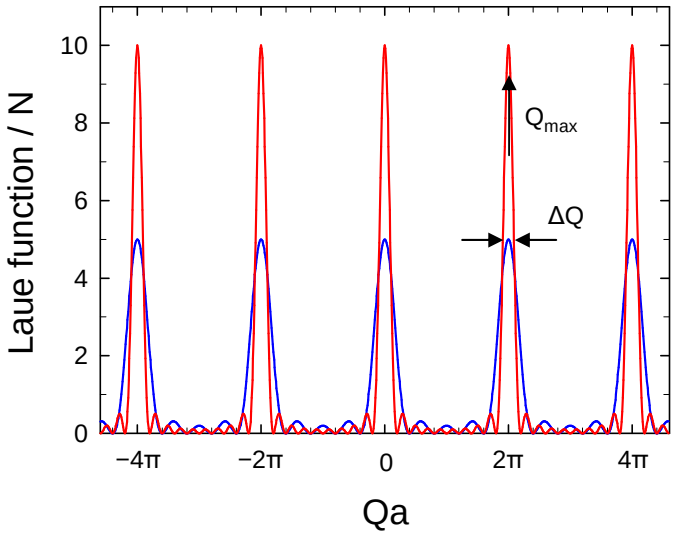


Fig. 3: *Laue function for a lattice with 5 (blue) or 10 (red) unit cells in the displayed Q-direction (normalised by $1/U$ analogous to the structure factor).*

that the main maxima occur if the arguments of the sines in the denominators are multiples of π. This implies $\mathbf{Q}_{\max} = (h2\pi/a_1, k2\pi/a_2, l2\pi/a_3)$ with all combinations of integer h, k, l (Miller indices). By using the formula for Q in the elastic case (4) we get

$$h\lambda = 2a_1 \sin\theta \tag{22}$$

and equivalent for the other spatial directions, the familiar Bragg conditions.
The value of the structure factor at these maxima is

$$S\left(\mathbf{Q}_{\max}\right) = \frac{1}{N} U^2 V^2 W^2 = N\,. \tag{23}$$

Thus, the height of the structure factor peaks (Bragg peaks) increases proportional to the number of scatterers. This is not trivial because the structure factor is normalised by $1/N$; so without any correlation it would not depend on the number of scatterers. Since we assume a perfect correlation here, we obtain this increase.
In order to obtain a large-N approximation for the width of the peaks, the slower-varying sine in the denominator can be linearised. It is sufficient to do this for the peak at $Q = 0$ because due to periodicity all peaks have the same shape. There,

$$\frac{\sin\left(UQ_x a_1/2\right)}{\sin\left(Q_x a_1/2\right)} \approx \frac{\sin\left(UQ_x a_1/2\right)}{Q_x a_1/2} = U j_0\left(UQ_x a_1/2\right)$$

where $j_0(x) = \sin x/x$ is the zeroth spherical Bessel function of the first kind. It can be seen that $S(\mathbf{Q})$ decays to half its maximum value in a Cartesian direction when this function decays to $1/\sqrt{2}$, i.e. at $x = 1.39\ldots$. Using the usual 'full width at half maximum' convention (see Fig. 3) one obtains for the width in x direction

$$\Delta Q_x = \frac{4 \cdot 1.39\ldots}{U a_1} = \frac{5.57\ldots}{U a_1}\,. \tag{24}$$

This is very close to the rule-of-thumb result $2\pi/\text{size}$ commonly found in the textbooks.

The preceding results are mostly valid also for general lattices where

$$\mathbf{r}_{uvw} = u\mathbf{a}_1 + v\mathbf{a}_2 + w\mathbf{a}_3 \tag{25}$$

with three arbitrarily oriented basis vectors $\mathbf{a}_{1,2,3}$. In the general case the condition that the sines have arguments of multiples of π is fulfilled for scattering vectors

$$\mathbf{Q}_{\max} = h\boldsymbol{\tau}_1 + k\boldsymbol{\tau}_2 + l\boldsymbol{\tau}_3 \tag{26}$$

where the *reciprocal lattice vectors* $\boldsymbol{\tau}_{1,2,3}$ are the solutions of the system of linear equations stating that $\mathbf{a}_\alpha \cdot \boldsymbol{\tau}_\alpha = 2\pi$ for $\alpha = 1, 2, 3$ and $\mathbf{a}_\alpha \cdot \boldsymbol{\tau}_\beta = 0$ for $\alpha \neq \beta$:

$$\begin{aligned}
\boldsymbol{\tau}_1 &= 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\,, \\
\boldsymbol{\tau}_2 &= 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\,, \\
\boldsymbol{\tau}_3 &= 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\,.
\end{aligned} \tag{27}$$

Finally, in real crystals there are often multiple scatterers in the unit cell (crystallographic basis) or there is a continuous distribution of scattering power within the unit cell. In these cases the total scattering density of the crystal is the convolution of the structure of the unit cell with that of the lattice discussed before. Therefore, the considerations of subsection 2.5 apply and the total scattering can be expressed as the product of the structure factor of the lattice (as before) and a form factor from the unit cell. Therefore, there is only a slight complication, a multiplication of the $S(\mathbf{Q})$ with a Q-dependent factor. Although the positions of the Bragg peaks are not affected by this multiplication, their intensity may change. Especially, Bragg peaks may vanish if they coincide with a zero of the Fourier transform of the unit cell. This situation regularly occurs for higher symmetries and is in detail covered by the reflection conditions tabulated e.g. in the International Tables for Crystallography [12].

2.4 Example: liquid structure factor

This second example is not an exact calculation but more a rough description of the features to be expected for scattering from a liquid (Fig. 4). A liquid also does not fulfil the requirement that the structure is static in the strict sense required above. Nevertheless, as will be derived in the section 3, a diffraction experiment will yield an $S(Q)$ corresponding to the instantaneous structure. But what is more a problem for the mathematical treatment is that for a given interparticle potential $V(r)$ there is no exact way to derive the pair correlation function $g(r)$. There are only approximative analytical methods [13] and numerical methods available for this purpose. Nevertheless it can be expected that there is a preferential nearest-neighbour distance r_{nn} which is roughly defined by the minimum of the interparticle potential and corresponds to a maximum in $g(r)$. As explained before $g(r)$ will drop sharply for too short distances because of the strong repulsion. For large r there will be no significant interaction between the particles so that the joint probability $\langle \rho(\mathbf{0})\rho(\mathbf{r})\rangle$ will become the product of the average densities ${\rho_0}^2$ and in consequence $\lim_{r\to\infty} g(r) = 1$.
From $g(r)$ by use of equation (17) the structure factor can be calculated. Although again an exact result cannot be given, several general features can be stated: For $Q \to \infty$, $\exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})$ becomes a rapidly oscillating function and the integral vanishes. Then one has

$$\lim_{Q\to\infty} S(Q) = 1\,. \tag{28}$$

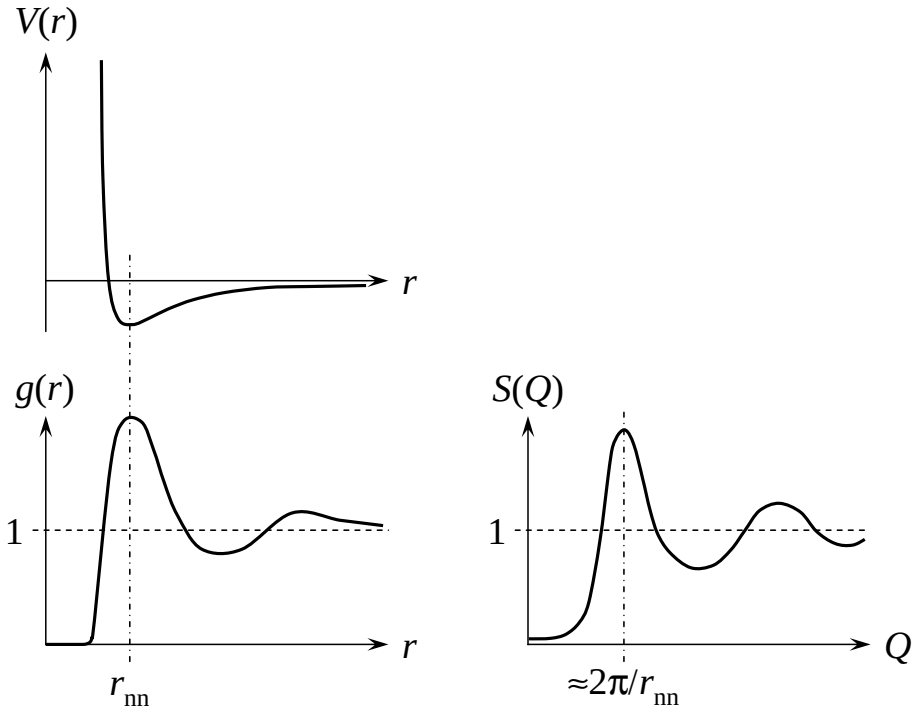


Fig. 4: *Schematic representation of interaction potential $V(r)$, pair correlation function $g(r)$, and scattering function $S(Q)$.*

For $Q \to 0$, $S(Q)$ measures only the overall density fluctuation, i.e. the fluctuation of the particle number:

$$\lim_{Q\to 0} S(Q) = \frac{V^2\langle\delta\rho^2\rangle}{N} = \frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle} = \rho_0 k_\mathrm{B} T \kappa_T \,. \tag{29}$$

Here, k_B denotes the Boltzmann factor, T the temperature and κ_T the isothermal compressibility. At intermediate Q, the structure factor of liquids shows a diminishing series of broad peaks, remainders of the Bragg peaks of a crystalline structure. The first peak occurs at a scattering vector roughly corresponding to the next neighbour distance by $Q_\mathrm{max} = 2\pi/r_\mathrm{nn}$.

2.5 Structure factor and form factor

Many physical systems are not constituted by point scatterers but nevertheless the density is not just some irregular function of coordinate. Rather, the scatterers are structured and this structure is identically repeated at different centres in space. Mathematically, this can be captured as a convolution:

$$\rho(\mathbf{r}) = \rho_\mathrm{centres}(\mathbf{r}) \otimes \rho_\mathrm{scatterer}(\mathbf{r}) = \int \mathrm{d}^3 r' \rho_\mathrm{centres}(\mathbf{r}')\rho_\mathrm{scatterer}(\mathbf{r}-\mathbf{r}') \tag{30}$$

where $\rho_\mathrm{centres}(\mathbf{r}) = \sum_j \delta(\mathbf{r} - \mathbf{r}_j^\mathrm{centres})$.
Now the convolution theorem (143) states that the Fourier transform of the density (30) is the product of the Fourier transforms of the underlying densities:

$$\mathcal{F}[\rho(\mathbf{r})] = \mathcal{F}[\rho_\mathrm{centres}(\mathbf{r})]\mathcal{F}[\rho_\mathrm{scatterer}(\mathbf{r})] \,. \tag{31}$$

Consequently, the scattering of the whole system, which for distinction should be denoted $I(Q)$ here, is

$$I(Q) = S(Q)P(Q) \tag{32}$$

where $S(Q)$ is the structure factor of the centre positions alone and $P(Q)$ is that of the density within a single scatterer[1]. The latter is usually called *form factor*.
The form factor is a (usually monotonically) decaying function, with a range of roughly $Q_{\max} \approx 2\pi/d$ where d is the size of the scatterer. In case of x-ray scattering from atoms d is in the Ångström range; therefore, $P(Q)$ decays within a couple of Å^{-1}. This implies that $S(Q)$ is damped significantly by (32) and higher order Bragg peaks in crystals or short range correlations in liquids cannot be observed well. For neutron scattering (nuclear, not magnetic) a potential with just a few femtometres range is relevant. There, the form factor would only decay at extremely high Q and is practically constant for all relevant Q values. Nevertheless, also for neutron scattering the form factor/strcture factor dichotomy is useful in the case of complex liquids (colloids, microemulsions, polymer solutions etc.). In that case, one has an arragement of statistically equal objects of nanometre size leading to a form factor decaying at a Q of the order of $0.1\ \text{Å}^{-1}$.
There is often some confusion in the use of "form factor" and "structure factor". The reason for this is that the two terms are defined relatively to each other. Imagine for example x-ray scattering from a polymer solution. The solution is constituted by polymer molecules which itself are constituted by atoms. Thus, there is a three-level hierarchy. With respect to the levels molecule-atom we have an atomic form factor $P_{\text{at}}(Q)$ and a polymer molecule structure factor $S_{\text{mol}}(Q)$. Looking at the levels solution-molecule, there is a molecular form factor $P_{\text{mol}}(Q)$ and a solution structure factor $S_{\text{sol}}(Q)$. But in the end $S_{\text{mol}}(Q) = P_{\text{mol}}(Q)$, meaning that the question whether the polymer molecule has a structure factor or a form factor depends on the view of this intermediate level.

3 Scattering from dynamic systems

Here, the more realistic situation will be considered in which the particles of the sample are moving. For moving particles, energy may be transferred to or from the scattered particle, or in the wave picture the frequency is changed by the Doppler effect. Thus an energy transfer

$$\Delta E = E' - E \equiv \hbar\omega \tag{33}$$

occurs, the scattering is in general *inelastic*, and $k' \neq k$. Q now does not anymore result from the isosceles construction in Fig. 2 drafted in black but from scattering triangles as those in blue and red. Application of the cosine theorem leads to the following expression for Q in the inelastic situation:

$$Q = \sqrt{k^2 + k'^2 - 2kk'\cos(2\theta)} \tag{34}$$

$$= \sqrt{\frac{8\pi^2}{\lambda^2} + \frac{2m\omega}{\hbar} - \frac{4\pi}{\lambda}\sqrt{\frac{4\pi^2}{\lambda^2} + \frac{2m\omega}{\hbar}}\cos(2\theta)} \tag{35}$$

[1]Note that in order to get equation (32) literally correct, one needs a different normalisation of $P(Q)$, namely by $1/N^2$ instead of $1/N$. This is done in different ways in the literature, so that formula (32) may appear with different additional prefactors, e.g. the average density within the particle.

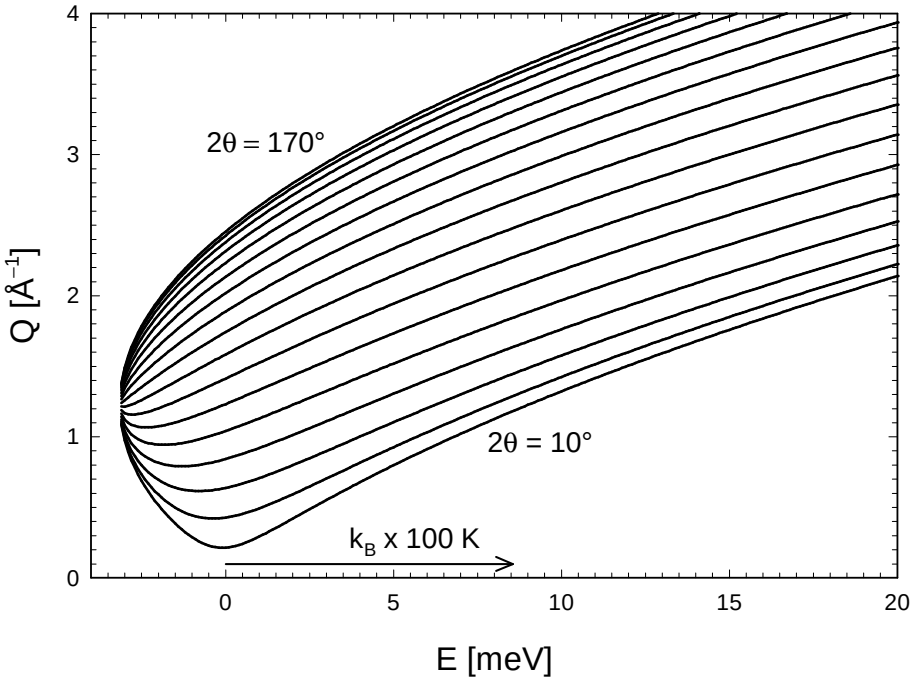


Fig. 5: *Scattering vectors Q accessed by a neutron scattering experiment with the detector at scattering angles $2\theta = 10 \ldots 170°$ vs. the energy transfer $\hbar\omega$ (incident wavelength $\lambda = 0.51$ nm). For comparison the thermal energy $k_\mathrm{B}T$ corresponding to 100 K is indicated by an arrow.*

for non-relativistic particles with rest mass $m > 0$, e.g. neutrons. Especially, it has to be observed now that Q also depends on $\hbar\omega$ implying that Q is not anymore constant for a single scattering angle. This complication may be unnecessary to consider in cases where $|\hbar\omega| \ll E$, as for x-ray scattering, Brillouin light scattering, and photon correlation spectroscopy. But it will be important for methods where $\hbar\omega$ and the incident energy of the particles E are comparable, as Raman light scattering or neutron scattering. Fig. 5 shows the magnitude of this effect for typical parameters of a neutron scattering experiment. It can be seen that it is by no means negligible for typical thermal energies of the sample even at temperatures as low as 100 K.
In analogy to (3) the *double differential cross-section* is defined as the probability density that a neutron is scattered into a solid angle element $\mathrm{d}\Omega$ with an energy transfer $\hbar\omega \ldots \hbar(\omega + \mathrm{d}\omega)$. A quantum-mechanical calculation based on Fermi's Golden Rule yields[2]:

$$\frac{\partial^2\sigma}{\partial\Omega\partial E'} \propto \frac{k'}{k} \sum_{\lambda,\lambda'} P_\lambda \left| \langle\lambda'|\langle\mathbf{k}'|\hat{V}|\mathbf{k}\rangle|\lambda\rangle \right|^2 \delta\left(\hbar\omega + E_\lambda - E_{\lambda'}\right) . \tag{36}$$

Here, $|\lambda\rangle$ and $|\lambda'\rangle$ are the states of the system before and after scattering, $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$ those of the scattered particle (plane waves). P_λ is the probability of the initial state, $\hat{V}$ is the interaction between scattered particle and system. Finally $\delta(\hbar\omega + E_\lambda - E_{\lambda'})$ expresses energy conservation. For different probes (neutrons, light, electrons etc.) and interaction mechanisms (e.g. nuclear, magnetic) $\hat{V}$ will be different and there may be additional variables characterising the state of

[2]The classical analogue of this formula can be easily derived [14] and is sufficient for certain applications, e.g. the dynamics of soft matter. But the classical approximation will fail for many of the systems discussed in this school. Therefore one needs the full quantum-mechanical apparatus. The derivation of (36) is explained in standard textbooks of neutron scattering [1–3] and will not be repeated here.

the scattered particle (spin, polarisation). Surprisingly, the most simple result emerges from nuclear neutron scattering mainly because the short-ranged nuclear potential can be replaced by the *Fermi pseudopotential*

$$\hat{V} \propto \sum_{j=1}^{N} b_j \delta(\mathbf{r} - \hat{\mathbf{r}}_j) \tag{37}$$

which is obviously the quantum-mechanical equivalent of the scattering length density (8). With this potential the inner matrix element can easily be calculated with the plane wave expressions for $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$:

$$\frac{\partial^2\sigma}{\partial\Omega\partial E'} = \frac{k'}{k} \sum_{\lambda} P_\lambda \sum_{\lambda'} \left| \sum_j \langle\lambda'| b_j \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j) |\lambda\rangle \right|^2 \delta\left(\hbar\omega + E_\lambda - E_{\lambda'}\right) . \tag{38}$$

Note that this expression neglects the spin of the neutron as well as that of the scattering system. Therefore, it is only valid for the situation of a polarised neutron beam in combination with spinless (or spin-polarised) sample nuclei.
For comparison, the expression for inelastic x-ray scattering is [15]:

$$\left.\frac{\partial^2\sigma}{\partial\Omega\partial E'}\right|_{\alpha\to\beta} \propto \frac{k'}{k} |{\boldsymbol{\epsilon}_\alpha}^* \cdot \boldsymbol{\epsilon}_\beta|^2 \sum_{\lambda} P_\lambda \sum_{\lambda'} \left| \sum_j \langle\lambda'| b_j \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j) |\lambda\rangle \right|^2 \delta\left(\hbar\omega + E_\lambda - E_{\lambda'}\right) . \tag{39}$$

It can be seen that the formula is complicated by the polarisation-dependent term $|{\boldsymbol{\epsilon}_\alpha}^* \cdot \boldsymbol{\epsilon}_\beta|^2$ but otherwise contains the same 'core terms' as (38). The same is true for all scattering probes. Therefore, we will continue with the simplest formulation in nuclear neutron scattering from spin zero nuclei.

3.1 Scattering functions

In order to related expression (36) to a correlation function one starts with an integral representation of the delta function (appendix: equation (137)):

$$\delta\left(\hbar\omega + E_\lambda - E_{\lambda'}\right) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \mathrm{d}t \exp\left(-\mathrm{i}\left(\omega + \frac{E_\lambda - E_{\lambda'}}{\hbar}\right) t\right) \tag{40}$$

which results from the fact that the delta function is the Fourier transform of a constant one. With this expression the matrix element in equation (36) can be written as a Fourier transform in time:

$$\begin{aligned}
& \left| \sum_j \langle\lambda'| b_j \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j) |\lambda\rangle \right|^2 \delta\left(\hbar\omega + E_\lambda - E_{\lambda'}\right) \\
= \; & \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \exp\left(-\mathrm{i}\frac{E_\lambda}{\hbar}t\right) \exp\left(-\mathrm{i}\frac{E_{\lambda'}}{\hbar}t\right) \\
& \qquad \sum_j b_j \langle\lambda'| \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j) |\lambda\rangle \sum_k b_k^* \langle\lambda| \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k) |\lambda'\rangle \\
= \; & \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \sum_{j,k} b_j b_k^* \langle\lambda| \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k) |\lambda'\rangle \\
& \qquad \langle\lambda'| \exp(\mathrm{i}E_{\lambda'}t/\hbar) \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j) \exp(-\mathrm{i}E_\lambda t/\hbar) |\lambda\rangle
\end{aligned} \tag{41}$$

If $\hat{\mathcal{H}}$ is the Hamiltonian of the scattering system, the fact that $|\lambda\rangle$ are energy eigenstates is expressed by

$$\hat{\mathcal{H}}|\lambda\rangle = E_\lambda|\lambda\rangle\,. \tag{42}$$

Iterating this equation n times yields:

$$\hat{\mathcal{H}}^n|\lambda\rangle = {E_\lambda}^n|\lambda\rangle\,. \tag{43}$$

By expanding the exponential into a power series one finally obtains from this relation

$$\exp(\mathrm{i}\hat{\mathcal{H}}t/\hbar)|\lambda\rangle = \exp(\mathrm{i}E_\lambda t/\hbar)|\lambda\rangle\,. \tag{44}$$

With this result and the analogous one for λ' it is possible to replace the eigenvalues E_λ in (41) by the Hamiltonian $\hat{\mathcal{H}}$:

$$\ldots\langle\lambda'|\exp(\mathrm{i}\hat{\mathcal{H}}t/\hbar)\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_i)\exp(-\mathrm{i}\hat{\mathcal{H}}t/\hbar)|\lambda\rangle\,. \tag{45}$$

In the picture of time dependent Heisenberg operators the application of the operator $\exp(\mathrm{i}\hat{\mathcal{H}}t/\hbar)$ and its conjugate just means a propagation by time t:

$$\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_i(t)\right) = \exp(\mathrm{i}\hat{\mathcal{H}}t/\hbar)\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_i(0))\exp(-\mathrm{i}\hat{\mathcal{H}}t/\hbar) \tag{46}$$

where we can arbitrarily set $\hat{\mathbf{r}}_i = \hat{\mathbf{r}}_i(0)$ because of translation of time invariance. Using this result the final expression for the double differential cross section is obtained:

$$\begin{aligned}\frac{\partial^2\sigma}{\partial\Omega\partial E'} &= \frac{k'}{k}\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t)\\ &\qquad\sum_\lambda P_\lambda\sum_{j,k}b_jb_k^*\left\langle\lambda\left|\exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(0))\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t))\right|\lambda\right\rangle\,.\end{aligned} \tag{47}$$

(Here, the sum over λ' vanishes because of the completeness relation $\sum_{\lambda'}|\lambda'\rangle\langle\lambda'| = 1$.) In addition the initial states of the scattering system are averaged weighted with the probability of their occurrence P_λ. The latter is given by the Boltzmann distribution

$$P_\lambda = \frac{1}{Z}\exp\left(-E_\lambda/k_\mathrm{B}T\right)\ \text{with}\ Z = \sum_\lambda\exp\left(-E_\lambda/k_\mathrm{B}T\right)\,. \tag{48}$$

As usual the thermal average is denoted by angular brackets, $\langle\ldots\rangle$. We now assume that the scatterers are identical with respect to the interactions ('chemically' identical) but may have different scattering lengths b_j but randomly distributed over the scatterers. In this case we have to do another averaging over all distributions of scattering lengths which is written as overline: $\overline{\ldots}$. Keeping in mind that for equal indices $b_ib_i^* = |b_i|^2$ has to be averaged while for unequal indices the scattering lengths itself will be averaged we end up with the usual separation into incoherent and coherent part:

$$\begin{aligned}\frac{\partial^2\sigma}{\partial\Omega\partial E'} &= \sum_\lambda P_\lambda\\ &\frac{k'}{k}\frac{\overline{|b|^2}-|\bar{b}|^2}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t)\sum_j\left\langle\lambda\left|\exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0))\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t))\right|\lambda\right\rangle\\ &+\frac{k'}{k}\frac{|\bar{b}|^2}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t)\sum_{j,k}\left\langle\lambda\left|\exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0))\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(t))\right|\lambda\right\rangle\,.\end{aligned} \tag{49}$$

The first term is the incoherent scattering. It involves the coordinate vector operators of the *same* atom at different times. The second, the coherent term correlates also *different* atoms at different times. The material dependent parts are now defined as the *scattering functions*[3]

$$S_{\mathrm{coh}}(Q,\omega) \equiv \frac{1}{2\pi N}\int_{-\infty}^{\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \sum_{j,k=1}^{N} \langle \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0)) \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(t))\rangle \,, \quad (50)$$

$$S_{\mathrm{inc}}(Q,\omega) \equiv \frac{1}{2\pi N}\int_{-\infty}^{\infty} \mathrm{d}t \exp(-\mathrm{i}\omega t) \sum_{j=1}^{N} \langle \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0)) \exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t))\rangle \,. \quad (51)$$

The former is called the *coherent* and the latter the *incoherent* scattering function. In terms of the scattering functions the double differential cross section can be written as

$$\frac{\partial^2\sigma}{\partial\Omega\partial E'} = \frac{1}{\hbar}\frac{k'}{k} N \left(\left(\overline{|b|^2} - |\bar{b}|^2 \right) S_{\mathrm{inc}}(\mathbf{Q},\omega) + |\bar{b}|^2 S_{\mathrm{coh}}(\mathbf{Q},\omega) \right) \,. \quad (52)$$

One can see that the incoherent term is weighted by the variance of scattering lengths. It did not show up in the considerations of section 2 because the assumption of equal scattering length excluded it. A treatment of the static case allowing a variance of the scattering length results in

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = N|\bar{b}|^2 S(\mathbf{Q}) + N\left(\overline{|b|^2} - |\bar{b}|^2 \right) \,. \quad (53)$$

Thus, the incoherent term exists also in scattering from a static system but it only constitutes a flat background.

The most common situation where incoherent scattering arises is that of neutron scattering. Because chemically identical atoms may be different isotopes a random variation of the scattering length is more the rule than the exception. It is also possible to 'smuggle in' the neglected dependence of neutron scattering on the spin orientation at that point. For neutron scattering with an unpolarised beam and without polarisation analysis, the random orientation of spins leads to a de-facto randomness of the scattering lengths having the same effect as a variation of isotopes[4].

Note also that the distinction between coherent and incoherent neutron scattering is often a bit blurred in case of polyatomic materials. Mostly the concept of partial structure factors (or partial scattering functions) is used, i.e. functions $S^{\alpha\beta}(\mathbf{Q})$ (or $S^{\alpha\beta}_{\mathrm{coh}}(\mathbf{Q},\omega)$) which represent the correlation of atoms of type α with those of type β. These have an incoherent counterpart $S^{\alpha}_{\mathrm{inc}}(\mathbf{Q},\omega)$ for those atoms α which show isotope or spin-disorder. On the other hand, if two elements are randomly exchangeable (e.g. in a mixed crystal) one could also describe this in the same way as isotope-incoherent scattering. In that case, even for x-rays one can speak about incoherent scattering because different elements have different electronic structures. A similar unclear terminology occurs in neutron scattering studies of polymers with different artificial replacements of hydrogen by deuterium.

Finally, there is often a discussion whether incoherent scattering may occur in light scattering. This is often demonstrated from polydisperse colloids. Because the form factor of a colloidal

[3]The definition of the scattering functions in literature usually differs by a factor $1/\hbar$ which is here included in equation (52). The difference is that here $S(Q,\omega)$ is a density in frequency (with unit [time]) while in the literature it is taken as a density in energy. As pointed out by J. Wuttke, the literature definition leads to an unnecessary complication of the sum rules in subsection 3.2.

[4]The inverse conclusion is true: Polarisation analysis in neutron scattering allows to separate coherent and spin-incoherent scattering.

particle depends on its size, this can be treated like a variation of the scattering length for nuclei in neutron scattering. But in the opinion of the author this does not constitute a case of incoherent scattering[5] because particles of different size also have different interaction, they are 'chemically different'. The only case where one can for certain speak of incoherent light scattering is that of particles which are made distinguishable by cores with different index of refraction [16].

3.2 Intermediate scattering functions

In some cases it is interesting to consider the part of expression (50) before the time-frequency Fourier transform, called *coherent intermediate scattering function*:

$$I_{\mathrm{coh}}(\mathbf{Q},t) = \frac{1}{N}\sum_{j,k=1}^{N}\left\langle \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0))\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(t))\right\rangle . \tag{54}$$

Its value for $t = 0$ expresses the correlation between atoms *at equal times*. On one hand equation (140) of the appendix tells that this is identical to the integral of the scattering function over all energy transfers:

$$I_{\mathrm{coh}}(\mathbf{Q},0) = \int_{-\infty}^{\infty} S_{\mathrm{coh}}(\mathbf{Q},\omega)\mathrm{d}\omega . \tag{55}$$

On the other hand this is just the (static) structure factor (10) because at equal t the position operators commute and the exponentials can be merged:

$$I_{\mathrm{coh}}(\mathbf{Q},0) = \frac{1}{N}\sum_{j,k=1}^{N}\left\langle \mathrm{e}^{\mathrm{i}\mathbf{Q}\cdot(\mathbf{r}_k-\mathbf{r}_j)}\right\rangle = S(\mathbf{Q}) . \tag{56}$$

Combining the two expressions for the intermediate coherent scattering function yields:

$$S(\mathbf{Q}) = \int_{-\infty}^{\infty} S_{\mathrm{coh}}(\mathbf{Q},\omega)\mathrm{d}\omega . \tag{57}$$

The concrete significance of this relation is that a diffraction experiment, which does not discriminate energies and thus implicitly integrates over all $\hbar\omega$, only shows the instantaneous correlation of the atoms, viz the structure of the sample[6]. $S(\mathbf{Q})$ is the structure factor as derived in section 2 for the static situation. The dynamic information is lost in the integration process.
Similarly the *incoherent intermediate scattering function* is

$$I_{\mathrm{inc}}(\mathbf{Q},t) = \frac{1}{N}\sum_{j=1}^{N}\left\langle \exp(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0))\exp(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t))\right\rangle \tag{58}$$

[5] ... in the sense of this lecture. Often the term "incoherent scattering" is used in the light scattering community as is "diffuse scattering" by neutron scatterers. In that sense, incoherent scattering would trivially exist in light scattering, except for ordered systems as opal or butterfly wings.

[6] Strictly speaking, this is only an approximation. There are several reasons why the integration in the diffraction experiment is not the 'mathematical' one of (57): (1) On the instrument the integral is taken along a curve of constant 2θ in Fig. 5 while constant Q would correspond to a horizontal line. (2) The double differential cross-section (52) contains a factor k'/k which depends on ω via (34). (3) The detector may have an efficiency depending on wavelength which will introduce another ω-dependent weight in the experimental integration. These effects can be taken into account in the so-called Placzek corrections [17, 18].

with the sum rule

$$I_{\mathrm{inc}}(\mathbf{Q},0)=\frac{1}{N}\sum_{j=1}^{N}\left\langle \mathrm{e}^{\mathrm{i}\mathbf{Q}\cdot(\mathbf{r}_j-\mathbf{r}_j)}\right\rangle = 1 = \int_{-\infty}^{\infty} S_{\mathrm{inc}}(\mathbf{Q},\omega)\mathrm{d}\omega\,. \tag{59}$$

Thus, the *incoherent* intermediate scattering function is normalised to one for each Q. Note that this result is independent of the actual structure of the sample. Also it is the same result as the high Q limit $S(Q)\rightarrow 1$ (28). This is a consequence of the more general fact that coherent and incoherent scattering become indistinguishable for large Q.

3.3 Van Hove correlation functions

As in the static situation, the scattering law can be traced back to distance distribution functions, the *van Hove correlation functions*. These can be derived by (inverse) Fourier transform of the intermediate scattering functions back into real space. In the coherent case:

$$G(\mathbf{r},t)=\left(\frac{1}{2\pi}\right)^3\int \mathrm{d}^3Q\exp\left(-\mathrm{i}\mathbf{Q}\cdot\mathbf{r}\right)\frac{1}{N}\sum_{j,k=1}^{N}\left\langle\exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0)\right)\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(t)\right)\right\rangle\,. \tag{60}$$

The derivation of the relation between the coherent dynamical structure factor $S_{\mathrm{coh}}(Q,\omega)$ and the generalized pair correlation function requires a strict quantum mechanical calculation. This problem results from the fact that the coordinate vector operators commute only at identical times. Therefore, in all algebraic manipulations the order of $\hat{\mathbf{r}}_j(0)$ and $\hat{\mathbf{r}}_j(t)$ must not be interchanged.
To begin, one writes the operator $\exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0)\right)$ as the Fourier transform of the delta function:

$$\exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_i(0)\right)=\int \mathrm{d}^3r'\delta\left(\mathbf{r}'-\hat{\mathbf{r}}_i(0)\right)\exp(-\mathrm{i}\mathbf{Q}\cdot\mathbf{r}')\,. \tag{61}$$

Using this expression equation (60) can be rewritten as

$$\begin{aligned} G(\mathbf{r},t) &= \left(\frac{1}{2\pi}\right)^3\frac{1}{N}\sum_{j,k=1}^{N}\Bigg\langle\int \mathrm{d}^3r'\delta\left(\mathbf{r}'-\hat{\mathbf{r}}_j(0)\right) \\ &\qquad \underbrace{\int \mathrm{d}^3Q\exp\left(-\mathrm{i}\mathbf{Q}\cdot\mathbf{r}-\mathrm{i}\mathbf{Q}\cdot\mathbf{r}'+\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_k(t)\right)}_{=(2\pi)^3\delta\left(\mathbf{r}+\mathbf{r}'-\hat{\mathbf{r}}_k(t)\right)}\Bigg\rangle \\ &= \frac{1}{N}\sum_{j,k=1}^{N}\int \mathrm{d}^3r'\left\langle\delta\left(\mathbf{r}-\mathbf{r}'+\hat{\mathbf{r}}_j(0)\right)\delta\left(\mathbf{r}'-\hat{\mathbf{r}}_k(t)\right)\right\rangle \end{aligned} \tag{62}$$

without changing the order of the operators at different times.
Now the particle density operator is introduced as a sum over delta functions at the particle position operators:

$$\hat{\rho}(\mathbf{r},t)\equiv\sum_{j=1}^{N}\delta\left(\mathbf{r}-\hat{\mathbf{r}}_j(t)\right)\,. \tag{63}$$

With this definition the pair correlation function can be written as time-dependent density-density correlation function:

$$G(\mathbf{r},t) = \frac{1}{N}\int \mathrm{d}^3r' \,\langle\hat{\rho}(\mathbf{r}'-\mathbf{r},0)\hat{\rho}(\mathbf{r}',t)\rangle \ . \tag{64}$$

With this form of the dynamic pair correlation function the dynamical structure factor can be—analogously to equation (16)—written as the double Fourier transform of the correlator of the particle density:

$$\begin{aligned} S_{\mathrm{coh}}(\mathbf{Q},\omega) &= \frac{1}{2\pi N}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t) && (65)\\ &\qquad \int \mathrm{d}^3r \int \mathrm{d}^3r' \exp\left(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}\right)\langle\hat{\rho}(\mathbf{r}'-\mathbf{r},0)\hat{\rho}(\mathbf{r}',t)\rangle && (66)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t)\int \mathrm{d}^3r \exp\left(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}\right) G(\mathbf{r},t)\,. && (67)\end{aligned}$$

Thus the scattering function is the double Fourier transform (in space and time) of the van Hove correlation function.
We now define the density operator in reciprocal space as the Fourier transform of (63):

$$\hat{\rho}_{\mathbf{Q}}(t) \equiv \sum_{j=1}^{N}\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t)\right) \tag{68}$$

and obtain for the dynamic structure factor

$$S_{\mathrm{coh}}(\mathbf{Q},\omega) = \frac{1}{2\pi N}\int_{-\infty}^{\infty}\mathrm{d}t\exp(-\mathrm{i}\omega t)\,\langle\hat{\rho}_{\mathbf{Q}}(0)\hat{\rho}_{-\mathbf{Q}}(t)\rangle \ . \tag{69}$$

Correspondingly, the intermediate scattering function is

$$I_{\mathrm{coh}}(\mathbf{Q},t) = \frac{1}{N}\,\langle\hat{\rho}_{\mathbf{Q}}(0)\hat{\rho}_{-\mathbf{Q}}(t)\rangle \tag{70}$$

which after insertion of (68) turns out to be equivalent to (54).
Analogously, one can define the van Hove *self* correlation function by setting $k = j$ in the preceding equations, leading to

$$G_{\mathrm{s}}(\mathbf{r},t) = \frac{1}{N}\sum_{j=1}^{N}\int \mathrm{d}^3r' \,\langle\delta\left(\mathbf{r}-\mathbf{r}'+\hat{\mathbf{r}}_j(0)\right)\delta\left(\mathbf{r}'-\hat{\mathbf{r}}_j(t)\right)\rangle \tag{71}$$

as the equivalent of (62). Note that there is no equivalent of equations (69) and (70) because the definition of the density already 'mixes' all particles and the product of the density with itself inevitably contains all pair correlations.
The pair correlation function has some general properties:

1. For spatially homogeneous systems the integrand in (64) is independent of $\mathbf{r}'$ which can be arbitrarily set to the origin $\mathbf{0}$:

$$G(\mathbf{r},t) = \frac{V}{N}\,\langle\hat{\rho}(-\mathbf{r},0)\hat{\rho}(\mathbf{0},t)\rangle = \frac{1}{\rho_0}\,\langle\hat{\rho}(\mathbf{0},0)\hat{\rho}(\mathbf{r},t)\rangle \ . \tag{72}$$

2. The pair correlation function has the following asymptotic behaviour: For fixed distance and $t \to \infty$ or fixed time and $r \to \infty$ the averages in equation (64) can be executed separately and in consequence

$$G(\mathbf{r},t) \to \frac{1}{N}\int \mathrm{d}^3r' \,\langle\hat{\rho}(\mathbf{r}'-\mathbf{r},0)\rangle\,\langle\hat{\rho}(\mathbf{r}',t)\rangle = \rho_0\,. \tag{73}$$

3. For $t = 0$ the operators commute and the convolution integral of equation (64) can be carried out:

$$G(\mathbf{r},0) = \frac{1}{N}\sum_{j,k=1}^{N}\langle\delta\left(\mathbf{r}+\hat{\mathbf{r}}_j(0)-\hat{\mathbf{r}}_k(0)\right)\rangle\ . \tag{74}$$

Comparison of this equation with (13) or putting $t = 0$ in (72) yields the relation to the static pair correlation function:

$$G(\mathbf{r},0) = \frac{1}{\rho_0}\langle\rho(\mathbf{0})\rho(\mathbf{r})\rangle \tag{75}$$

or in terms of the commonly defined $g(r)$:

$$G(\mathbf{r},0) = \delta(\mathbf{r}) + \rho_0 g(\mathbf{r})\,. \tag{76}$$

This equation expresses again the fact that the diffraction experiment ($g(r)$) gives an average snapshot picture ($G(r,0)$) of the sample.

3.4 Classical approximation

In the classical approximation the operators can be replaced by variables. Especially the position operators reduce to trajectories of particles, $\mathbf{r}_j(t)$. Then the integrals of equations (62) and (71) can be carried out and yield

$$G^{\mathrm{cl}}(\mathbf{r},t) = \frac{1}{N}\sum_{j,k=1}^{N}\delta(\mathbf{r}-\mathbf{r}_k(t)+\mathbf{r}_j(0)) \text{ and} \tag{77}$$

$$G_{\mathrm{s}}^{\mathrm{cl}}(\mathbf{r},t) = \frac{1}{N}\sum_{j=1}^{N}\delta(\mathbf{r}-\mathbf{r}_j(t)+\mathbf{r}_j(0))\,, \tag{78}$$

respectively. The former equation expresses the probability to find *any* particle at a time t in a distance $\mathbf{r}$ from another at time 0. The latter equation denotes this probability for the *same* particle. It therefore depends only on the particle's displacement during a time interval $\Delta\mathbf{r}_j(t) = \mathbf{r}_j(t) - \mathbf{r}_j(0)$ leading to a simple expression for the intermediate incoherent scattering function:

$$I_{\mathrm{inc}}^{\mathrm{cl}}(\mathbf{Q},t) = \frac{1}{N}\sum_{j=1}^{N}\langle\exp\left(-\mathrm{i}\mathbf{Q}\cdot\Delta\mathbf{r}_j(t)\right)\rangle\ . \tag{79}$$

In certain cases (if $G_{\mathrm{s}}^{\mathrm{cl}}(\mathbf{r},t)$ is a Gaussian in space) this expression can be further simplified using the "Gaussian approximation":

$$I_{\mathrm{inc}}^{\mathrm{Gauss}}(Q,t) = \exp\left(-\frac{1}{6}Q^2\langle\Delta r^2(t)\rangle\right)\ . \tag{80}$$

Here $\langle \Delta r^2(t) \rangle$ is the average mean squared displacement which often follows simple laws, e.g. $\langle \Delta r^2(t) \rangle = 6Dt$ for simple diffusion. Because one of the prerequisites of the Gaussian approximation is that all particles move statistically in the same way (dynamic homogeneity) the particle average and the index i vanish. An analogous expression can be derived for the coherent scattering.

The most prominent difference between the correct quantum-mechanical treatment and the classical is that the quantum-mechanical scattering functions are asymmetric with respect to the energy transfer $\hbar\omega$,

$$S_{[\mathrm{coh|inc}]}(\mathbf{Q}, -\omega) = \exp\left(\frac{\hbar\omega}{k_\mathrm{B}T}\right) S_{[\mathrm{coh|inc}]}(\mathbf{Q}, \omega)\,, \tag{81}$$

while those in classical approximation are symmetric:

$$S^{\mathrm{cl}}_{[\mathrm{coh|inc}]}(\mathbf{Q}, -\omega) = S^{\mathrm{cl}}_{[\mathrm{coh|inc}]}(\mathbf{Q}, \omega)\,. \tag{82}$$

Equation (81) expresses the fact that the probability for a neutron to be scattered with energy loss is always higher than the probability to be scattered with energy gain (Fig. 6). This can be understood as a *detailed balance* factor: In equilibrium, the probability for the scattering system to be in the lower energy state is higher by the factor $\exp(\hbar\omega/k_\mathrm{B}T)$. Therefore the probability of scattering into a state with higher energy is more probable by the same factor than scattering into the lower energy state. The effect of the asymmetry will be noticeable for low temperatures and high energy transfers unless $T \gg \hbar\omega/k_\mathrm{B}$ [7]. Considering that for room temperature $k_\mathrm{B} \cdot 300\,\mathrm{K} \approx 26$ meV it is clear that the condition for a classical treatment is often not fulfilled in neutron scattering.

From the asymmetry of $S(\mathbf{Q}, \omega)$ follows that the intermediate scattering function $I(\mathbf{Q}, t)$ and the van Hove correlation function $G(\mathbf{r}, t)$ are *complex*. This is surprising but allowed because they are (in contrast to within the classical treatment) no observable quantities[8] as $S(\mathbf{Q}, \omega)$ which still has to be real. By inversion of the Fourier transform in time it follows for the intermediate scattering functions and the van Hove correlation functions:

$$I_{[\mathrm{coh|inc}]}(\mathbf{Q}, -t) = I_{[\mathrm{coh|inc}]}\left(\mathbf{Q}, t - \frac{\mathrm{i}\hbar}{k_\mathrm{B}T}\right) \tag{83}$$

$$G_{[\mathrm{s}]}(\mathbf{r}, -t) = G_{[\mathrm{s}]}\left(\mathbf{r}, t - \frac{\mathrm{i}\hbar}{k_\mathrm{B}T}\right) \tag{84}$$

The latter relations follow from the quantum-mechanical peculiarity that time-dependent operators in a correlation function may not be interchanged but $\langle \hat{A}(0)\hat{B}(t) \rangle = \langle \hat{B}(t - \mathrm{i}\hbar/k_\mathrm{B}T)\hat{A}(0) \rangle$ and $\langle \hat{A}(0)\hat{B}(t) \rangle^* = \langle \hat{B}^\dagger(t)\hat{A}^\dagger(0) \rangle$.

In many cases, the classical particle trajectories are much easier to derive that their quantum-mechanical counterparts. If a full quantum-mechanical result is not available, at least an approximation to the correct scattering functions can be derived [19] which fulfills the detailed balance

[7] There is another condition for the validity of the classical calculation concerning the momentum transfer, $Q \ll \sqrt{2Mk_\mathrm{B}T}/\hbar$, where M is the mass of the scattering particle. The meaning of this condition is that the De Broglie wavelength of the scatterer should be sufficiently small compared to the length scale of the scattering experiment $1/Q$.

[8] Note that the neutron spin echo spectrometer, which is said to measure the intermediate scattering function, performs an inverse *cosine* Fourier transform instead of an exponential one. Therefore it actually measures only the real part of the intermediate scattering function.

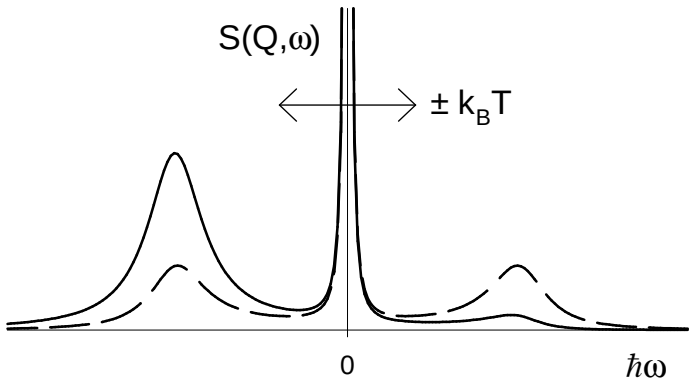


Fig. 6: *Example of the asymmetry of the scattering function due to the detailed balance factor. The (neutron) energy loss side is enhanced compared to the energy gain side for the correct quantum-mechanical result (continuous curve). The classical result (dashed curve) in contrast is mirror symmetric. The arrow indicates for comparison the thermal energy* $\pm k_{\mathrm{B}}T$.

relation (81). It is based on simply multiplying the square root of the prefactor in equation (81) neutron energy loss side and dividing it out of the energy gain side:

$$S_{[\mathrm{coh|inc}]}(\mathbf{Q},\omega) \approx \exp\left(-\frac{\hbar\omega}{2k_{\mathrm{B}}T}\right) S^{\mathrm{cl}}_{[\mathrm{coh|inc}]}(\mathbf{Q},\omega)\,. \tag{85}$$

The corresponding result for the intermediate scattering functions and the van Hove correlation functions results from shifting by half the imaginary time of expressions (83) and (84):

$$I_{[\mathrm{coh|inc}]}(\mathbf{Q},t) \approx I^{\mathrm{cl}}_{[\mathrm{coh|inc}]}\left(\mathbf{Q}, t+\frac{\mathrm{i}\hbar}{2k_{\mathrm{B}}T}\right) \tag{86}$$

$$G_{[\mathrm{s}]}(\mathbf{r},t) \approx G^{\mathrm{cl}}_{[\mathrm{s}]}\left(\mathbf{r}, t+\frac{\mathrm{i}\hbar}{2k_{\mathrm{B}}T}\right)\,. \tag{87}$$

3.5 Example: ideal gas

Although this is possibly the simplest system one can imagine, the calculation is already rather intricate. The simplicity of the example is mainly based on the definition of an ideal gas, that particles do not interact. This implies that there are no correlations between different particles and the $j \neq k$ terms vanish in expressions as (50). Therefore, the incoherent and coherent quantities are equal for the ideal gas:

$$S_{\mathrm{coh}}(Q,\omega) = S_{\mathrm{inc}}(Q,\omega)\,,\; I_{\mathrm{coh}}(Q,t) = I_{\mathrm{inc}}(Q,t)\,,\; G(r,t) = G_{\mathrm{s}}(r,t)\,. \tag{88}$$

(Because the ideal gas is isotropic, r and Q are scalars.) In addition, all particles behave statistically in the same way and therefore the averages over all particles can be replaced by a single representative particle:

$$\begin{aligned} I(Q,t) &= \frac{1}{N}\sum_{j=1}^{N}\left\langle \exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(0)\right)\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_j(t)\right)\right\rangle \\ &= \left\langle \exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1(0)\right)\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1(t)\right)\right\rangle\,. \end{aligned} \tag{89}$$

Writing out the Heisenberg operator $\hat{\mathbf{r}}_1(t)$ in its explicit form one obtains

$$I(Q,t) = \left\langle \exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1(0)\right) \exp\left(\mathrm{i}\hat{\mathcal{H}}_1 t/\hbar\right) \exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1(t)\right) \exp\left(-\mathrm{i}\hat{\mathcal{H}}_1 t/\hbar\right) \right\rangle \tag{90}$$

where $\hat{\mathcal{H}}_1$ is the Hamiltonian of the representative particle which is simply the square of the momentum operator divided by twice the scattering particle's mass:

$$\hat{\mathcal{H}}_1 = \frac{1}{2M}\hat{\mathbf{p}}^2 \,. \tag{91}$$

Taking into account that the operator $\exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1\right)$ shifts the momentum

$$\exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1\right)\hat{\mathbf{p}}\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}_1\right) = \hat{\mathbf{p}} + \hbar\mathbf{Q} \tag{92}$$

one gets

$$I(Q,t) = \left\langle \exp\left(\mathrm{i}\hat{\mathcal{H}}_1' t/\hbar\right) \exp\left(-\mathrm{i}\hat{\mathcal{H}}_1 t/\hbar\right) \right\rangle \tag{93}$$

where $\hat{\mathcal{H}}_1'$ denotes the single-particle Hamiltonian with shifted momentum:

$$\hat{\mathcal{H}}_1' = \frac{1}{2M}\left(\hat{\mathbf{p}} + \hbar\mathbf{Q}\right)^2 = \hat{\mathcal{H}}_1 + \frac{\hbar}{M}\mathbf{Q}\cdot\hat{\mathbf{p}} + \frac{\hbar^2 Q^2}{2M} \,. \tag{94}$$

Insertion of (94) into (93) yields:

$$I(Q,t) = \exp\left(\frac{\mathrm{i}\hbar t Q^2}{2M}\right) \left\langle \exp\left(\frac{\mathrm{i}t\mathbf{Q}\cdot\hat{\mathbf{p}}}{M}\right) \right\rangle \tag{95}$$

The thermodynamic average in this expression can be calculated with the equilibrium distribution of momenta. Here, a Boltzmann distribution is assumed:

$$\left\langle \exp\left(\mathrm{i}t\mathbf{Q}\cdot\hat{\mathbf{p}}\right)\right\rangle = \frac{\int \mathrm{d}^3p \, \exp\left(-\frac{p^2}{2Mk_\mathrm{B}T}\right) \exp\left(\mathrm{i}t\mathbf{Q}\cdot\mathbf{p}\right)}{\int \mathrm{d}^3p \, \exp\left(-\frac{p^2}{2Mk_\mathrm{B}T}\right)} = \exp\left(-Q^2t^2k_\mathrm{B}T/2M\right) \,. \tag{96}$$

From (95) and (96) the intermediate scattering function is finally obtained:

$$I(Q,t) = \exp\left(-\frac{Q^2}{2M}\left(k_\mathrm{B}Tt^2 + \mathrm{i}\hbar t\right)\right) \,. \tag{97}$$

In the same way as the Fourier transform of the Gaussian (153), the Fourier transform of (97) can be calculated yielding

$$S(Q,\omega) = \sqrt{\frac{M}{2\pi k_\mathrm{B}TQ^2}} \exp\left(-\frac{M}{2\pi k_\mathrm{B}TQ^2}\left(\omega + \frac{\hbar Q^2}{2M}\right)^2\right) \,. \tag{98}$$

The scattering function is a Gaussian distribution around $-\hbar Q^2/2M$ (Fig. 7) showing that on average the neutrons lose the 'recoil energy' $E_\mathrm{r} = \hbar^2Q^2/2M$ during the scattering event. The width of the Gaussian, $\sqrt{k_\mathrm{B}T/M}\hbar Q$ increases with temperature and scattering 'vector' Q.
The van Hove correlation function can be calculated immediately by inverse Fourier transform from (97) because $I(Q,t)$ is also a Gaussian in Q:

$$G(r,t) = \left(\frac{M}{2\pi k_\mathrm{B}Tt(t + \mathrm{i}\hbar/k_\mathrm{B}T)}\right)^{3/2} \exp\left(-\frac{Mr^2}{2k_\mathrm{B}Tt(t + \mathrm{i}\hbar/k_\mathrm{B}T)}\right) \,. \tag{99}$$

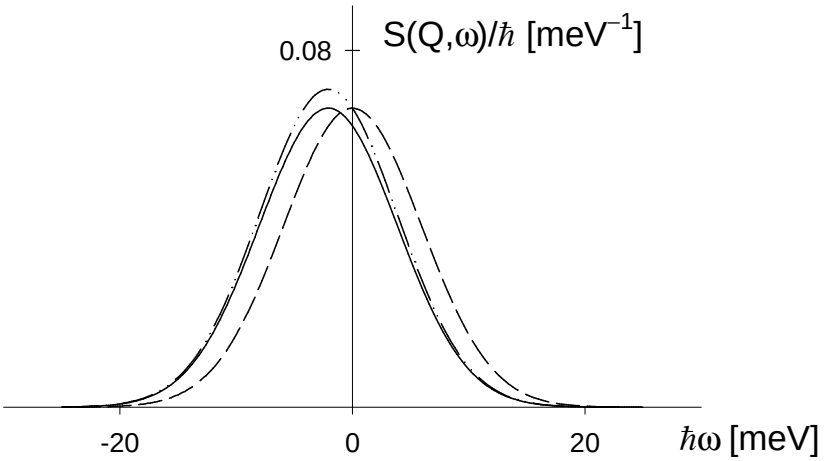


Fig. 7: *Scattering from an ideal gas calculated with the parameters of helium at 100 K for $Q = 2$ Å^{-1}. The continuous curve shows the correct quantum-mechanical result, the dashed curve that of a classical calculation. The dot-dashed curve represents the approximation resulting from applying (85) to the classical result.*

To demonstrate the differences arising from a classical calculation, the classical intermediate scattering function will be derived too. Because of the identity of the particles (58) reduces to

$$I^{\mathrm{cl}}(Q,t) = \langle \exp\left(\mathrm{i}\mathbf{Q}\cdot(\mathbf{r}_1(t) - \mathbf{r}_1(0))\right)\rangle \ . \tag{100}$$

In an ideal gas the trajectory of a particle is $\mathbf{r}_1(t) = \mathbf{r}_1(0) + \mathbf{v}t$ yielding:

$$I^{\mathrm{cl}}(Q,t) = \langle \exp\left(\mathrm{i}\mathbf{Q}\cdot\mathbf{v}t\right)\rangle \ . \tag{101}$$

This thermodynamic average can be calculated using the Maxwell distribution of velocities:

$$P(v) = \sqrt{\frac{2\pi k_{\mathrm{B}}T}{M}}\exp\left(-\frac{Mv^2}{2k_{\mathrm{B}}T}\right) \tag{102}$$

resulting in

$$I^{\mathrm{cl}}(Q,t) = \exp\left(-\frac{k_{\mathrm{B}}TQ^2t^2}{2M}\right) \ . \tag{103}$$

In contrast to the quantum-mechanical $I(Q,t)$ (97), this quantity is real and (accidentally for the ideal gas) $= |I(Q,t)|$. Because $I^{\mathrm{cl}}(Q,t)$ is a Gaussian in both Q and r, the Fourier transforms can be calculated by direct application of (153). The scattering function is

$$S^{\mathrm{cl}}(Q,\omega) = \sqrt{\frac{M}{2\pi k_{\mathrm{B}}TQ^2}}\exp\left(-\frac{M\omega^2}{2\pi k_{\mathrm{B}}TQ^2}\right) \ . \tag{104}$$

a Gaussian of the same width as the quantum-mechanical result (98) but centred around zero energy transfer (Fig. 7). Thus, classically neutrons are scattered with no average energy transfer. The classical van Hove correlation function

$$G^{\mathrm{cl}}(r,t) = \left(\frac{m_{\mathrm{sc}}}{2\pi k_{\mathrm{B}}Tt^2}\right)^{3/2}\exp\left(-\frac{m_{\mathrm{sc}}r^2}{2k_{\mathrm{B}}Tt^2}\right) \tag{105}$$

conveys the meaning that at any time the distribution of distances travelled by particles of an ideal gas is Gaussian with a width increasing linearly in time.
Finally, the approximations of the correct quantum-mechanical results from the classical using (86) and (85) are

$$\tilde{I}(Q,t) = \exp\left(-\frac{k_\mathrm{B}TQ^2}{2M}\left(t+\frac{\mathrm{i}\hbar}{2k_\mathrm{B}T}\right)^2\right) \quad \text{and} \tag{106}$$

$$\tilde{S}(Q,\omega) = \sqrt{\frac{M}{2\pi k_\mathrm{B}TQ^2}}\exp\left(-\frac{M}{2\pi k_\mathrm{B}TQ^2}\left(\omega^2+\frac{\hbar Q^2}{M}\omega\right)\right) . \tag{107}$$

It can be seen (Fig. 7) that the approximation captures the shift by the recoil energy correctly but the normalisation is wrong: $\int \tilde{S}(Q,\omega)\mathrm{d}\omega = \tilde{I}(Q,0) = \exp(\hbar^2Q^2/8Mk_\mathrm{B}T) \neq 1$. Nevertheless, equations (106) and (107) are different from (97) and (98) only in the order $\hbar^2$, which makes them better approximations than the purely classical results deviating already in $\hbar^1$ terms.

3.6 Example: harmonic oscillator

In this example we will deal with the inelastic scattering of a single particle in a harmonic potential. Although this example does not have a direct physical relevance, it already shows the basic properties of phonon scattering which will be discussed in lecture B4 in more detail.
At first, some elementary quantum-mechanical results on the harmonic oscillator will be recapitulated. As shown later, we only need these in one dimension even if the actual potential in three-dimensional. The Hamiltonian is then

$$\hat{\mathcal{H}} = \frac{1}{2M}\hat{p}^2 + \frac{M{\omega_0}^2}{2}\hat{x}^2 . \tag{108}$$

The first term is the kinetic energy and the second one the potential energy where the force constant is already replaced by the resonance frequency: $\omega = \sqrt{k/M}$. The Hamiltonian can be written in a simple way by using the Bose operators

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{M\omega_0}{\hbar}}\hat{x} + \mathrm{i}\sqrt{\frac{1}{M\hbar\omega_0}}\hat{p}\right) , \tag{109}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{M\omega_0}{\hbar}}\hat{x} - \mathrm{i}\sqrt{\frac{1}{M\hbar\omega_0}}\hat{p}\right) \tag{110}$$

as

$$\hat{\mathcal{H}} = \hbar\omega_0\left(\hat{a}\hat{a}^\dagger + 1/2\right) . \tag{111}$$

The eigenvalues of this equation are

$$E_n = (n+1/2)\hbar\omega_0 \tag{112}$$

and the eigenfunctions fulfill the relations

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle , \tag{113}$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle , \tag{114}$$

$$\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle . \tag{115}$$

Thus, the operators $\hat{a}^\dagger$ and $\hat{a}$ correspond to the phonon creation and annihilation operators. The equation of motion of the harmonic oscillator in Heisenberg representation (which is needed to calculate the scattering function) is:

$$\mathrm{i}\hbar\frac{\partial\hat{a}}{\partial t} = \left[\hat{a}, \hat{\mathcal{H}}\right] = \hbar\omega_0\hat{a}\,. \tag{116}$$

The formal solutions for the Bose operators are

$$\hat{a}(t) = \hat{a}(0)\exp(-\mathrm{i}\omega_0 t)\,, \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)\exp(\mathrm{i}\omega_0 t)\,, \tag{117}$$

and with this the time-dependent position operator is

$$\hat{x}(t) = \sqrt{\frac{\hbar}{2M\omega_0}}\left(\hat{a}(0)\exp(-\mathrm{i}\omega_0 t) + \hat{a}^\dagger(0)\exp(\mathrm{i}\omega_0 t)\right)\,. \tag{118}$$

As in the case of the ideal gas coherent and incoherent scattering are the same because we consider only a single particle:

$$I_{\mathrm{coh}}(\mathbf{Q},t) = I_{\mathrm{inc}}(\mathbf{Q},t) = \left\langle \exp\left(-\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}(0)\right)\exp\left(\mathrm{i}\mathbf{Q}\cdot\hat{\mathbf{r}}(t)\right)\right\rangle\,. \tag{119}$$

By an intricate calculation which is carried out in chapters 3.5, 3.6 and 3.8 of ref. 3 it is possible to separate the argument of the exponential into a time-independent and a time-dependent correlation function:

$$I(\mathbf{Q},t) = \exp\left(-\left\langle(\mathbf{Q}\cdot\hat{\mathbf{r}})^2\right\rangle + \left\langle(\mathbf{Q}\cdot\hat{\mathbf{r}}(0))\,(\mathbf{Q}\cdot\hat{\mathbf{r}}(t))\right\rangle\right)\,. \tag{120}$$

(Note that this calculation is not based on a general theorem but requires that $\hat{x}$ is a linear function of the Bose operators.) Because the first correlator is just the second one at $t = 0$, it is sufficient to calculate the latter: By defining the x direction parallel to $\mathbf{Q}$ one gets

$$\left\langle(\mathbf{Q}\cdot\hat{\mathbf{r}}(0))\,(\mathbf{Q}\cdot\hat{\mathbf{r}}(t))\right\rangle = Q^2\langle\hat{x}(0)\hat{x}(t)\rangle\,.$$

Inserting the solution (118) one continues:

$$= \frac{\hbar Q^2}{2M\omega_0}\left(\left(1+\left\langle\hat{a}^\dagger\hat{a}\right\rangle\right)\exp(\mathrm{i}\omega_0 t) + \left\langle\hat{a}^\dagger\hat{a}\right\rangle\exp(-\mathrm{i}\omega_0 t)\right)$$

Now, $\langle\hat{a}^\dagger\hat{a}\rangle$ is the average harmonic's index (corresponding to the phonon count) following the Bose statistics: $\langle n\rangle = 1/(\exp(\hbar\omega_0/k_{\mathrm{B}}T) - 1)$:

$$\begin{aligned}
&= \frac{\hbar Q^2}{2M\omega_0}\left(\frac{\exp(\hbar\omega_0/k_{\mathrm{B}}T)}{\exp(\hbar\omega_0/k_{\mathrm{B}}T)-1}\exp(\mathrm{i}\omega_0 t) + \frac{1}{\exp(\hbar\omega_0/k_{\mathrm{B}}T)-1}\exp(-\mathrm{i}\omega_0 t)\right)\\
&= \frac{\hbar Q^2}{2M\omega_0}\left(\frac{\exp(\hbar\omega_0/2k_{\mathrm{B}}T)\exp(\mathrm{i}\omega_0 t)}{\exp(\hbar\omega_0/2k_{\mathrm{B}}T) - \exp(-\hbar\omega_0/2k_{\mathrm{B}}T)}\right.\\
&\qquad\left. + \frac{\exp(-\hbar\omega_0/2k_{\mathrm{B}}T)\exp(-\mathrm{i}\omega_0 t)}{\exp(\hbar\omega_0/2k_{\mathrm{B}}T) - \exp(-\hbar\omega_0/2k_{\mathrm{B}}T)}\right)
\end{aligned}$$

So the final result is

$$\left\langle(\mathbf{Q}\cdot\hat{\mathbf{r}}(0))\,(\mathbf{Q}\cdot\hat{\mathbf{r}}(t))\right\rangle = \frac{\hbar Q^2}{2M\omega_0}\frac{\cosh(\hbar\omega_0/2k_{\mathrm{B}}T + \mathrm{i}\omega_0 t)}{\sinh(\hbar\omega_0/2k_{\mathrm{B}}T)} \tag{121}$$

which for $t = 0$ simplifies to

$$\left\langle (\mathbf{Q} \cdot \hat{\mathbf{r}})^2 \right\rangle = \frac{\hbar Q^2}{2M\omega_0} \coth\left(\hbar\omega_0 / 2k_\mathrm{B}T\right) \tag{122}$$

Inserting (121) and (122) into (120) yields the intermediate scattering function of the harmonic oscillator,

$$I(Q,t) = \exp\left(-\frac{\hbar Q^2}{2M\omega_0} \coth\left(\hbar\omega_0/2k_\mathrm{B}T\right)\right) \exp\left(\frac{\hbar Q^2}{2M\omega_0} \frac{\cosh\left(\hbar\omega_0/2k_\mathrm{B}T + \mathrm{i}\omega_0 t\right)}{\sinh\left(\hbar\omega_0/2k_\mathrm{B}T\right)}\right) \tag{123}$$

which, as expected from symmetry, only depends on the magnitude of $\mathbf{Q}$, not its direction.
The Fourier transform of the intermediate scattering function into frequency domain (energy transfer of the inelastic scattering) is possible by a series expansion in terms of modified Bessel functions, $\exp(y\cosh x) = \sum_{m=-\infty}^{\infty} I_m(y)\exp(mx)$ [20]. With this, one gets

$$\begin{aligned} I(Q,t) &= \exp\left(-\frac{\hbar Q^2}{2M\omega_0} \coth\left(\hbar\omega_0/2k_\mathrm{B}T\right)\right) \\ &\quad \times \sum_{m=-\infty}^{\infty} I_m\left(\frac{\hbar Q^2}{2M\omega_0} \mathrm{csch}\ \left(\hbar\omega_0/2k_\mathrm{B}T\right)\right) \exp\left(m\omega_0\left(\mathrm{i}t + \frac{\hbar}{2k_\mathrm{B}T}\right)\right) \end{aligned}$$

($\mathrm{csch}\, x \equiv 1/\sinh x$ is the hyperbolic cosecant) where every term contains a complex exponential oscillating with frequency $m\omega_0$. Each of these transform into a delta function. Therefore, the scattering function of the harmonic oscillator is an infinite sum of delta functions with weights depending on the energy transfer:

$$\begin{aligned} S(Q,\omega) &= \underbrace{\exp\left(-\frac{\hbar Q^2}{2M\omega_0} \coth\left(\hbar\omega_0/2k_\mathrm{B}T\right)\right)}_{\text{Debye-Waller factor}} \underbrace{\exp\left(-\frac{\hbar\omega}{2k_\mathrm{B}T}\right)}_{\text{detailed balance factor}} \\ &\quad \times \sum_{m=-\infty}^{\infty} I_m\left(\frac{\hbar Q^2}{2M\omega_0} \mathrm{csch}\ \left(\hbar\omega_0/2k_\mathrm{B}T\right)\right) \delta(\omega - m\omega_0)\,. \end{aligned} \tag{124}$$

As can be seen, the prefactor have an immediate physical significance. It is also possible to explain the terms of the sum in the picture of phonons:

$$\begin{aligned} S(Q,\omega) = \textit{DWF} \times \textit{DBF} \times \big(\quad & \delta(\omega) && \text{elastic scattering} \\ & + \ldots \delta(\omega - \omega_0) && \text{one phonon creation} \\ & + \ldots \delta(\omega + \omega_0) && \text{one phonon annihilation} \\ & + \ldots \delta(\omega - 2\omega_0) && \text{two phonon creation} \\ & + \ldots \delta(\omega + 2\omega_0) && \text{two phonon annihilation} \\ & + \ldots \quad \big) \end{aligned}$$

Appendices

A Elementary properties of the Fourier transform

In this section only those properties of the Fourier transform will be recapitulated which are relevant for the understanding of the results on scattering presented in the lecture[9]. The proofs

[9]For a more in-depth introduction into the properties of Fourier transforms relevant for scattering see lecture I of the JCNS Neutron Scattering Laboratory Course 2008 [21].

of the theorems will mostly only be sketched. For the mathematically more inclined reader refs. 22–24 are recommended.
According to the dominant standard in neutron scattering the Fourier transform will be defined as

$$F(X) = \int_{-\infty}^{\infty} f(x) \exp(\mathrm{i}Xx)\mathrm{d}x\,. \tag{125}$$

Wherever possible I will use the functional notation

$$\mathcal{F}_{X|\,x}[f(x)] = \int_{-\infty}^{\infty} f(x) \exp(\mathrm{i}Xx)\mathrm{d}x\,. \tag{126}$$

to abbreviate the Fourier integral. The inversion of the FT (125) is:

$$f(x) = \mathcal{F}^{-1}_{x|\,X}[F(X)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(X) \exp(-\mathrm{i}Xx)\mathrm{d}X\,. \tag{127}$$

The proof that (127) is the inverse of (125) is not simple and can be found in ref. 22.
A straightforward generalisation of the FT is that to multiple dimensions. The most important case of three dimensions is:

$$F(X,Y,Z) = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{\infty}\mathrm{d}z \exp(\mathrm{i}Xx)\exp(\mathrm{i}Yy)\exp(\mathrm{i}Zz)f(x,y,z)\,. \tag{128}$$

The arguments of the exponentials can be grouped together and replaced by the scalar product of the vector $\mathbf{r} = (x,y,z)$ and the 'reciprocal space' vector $\mathbf{Q} = (X,Y,Z)$:

$$F(\mathbf{Q}) = \int \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})f(\mathbf{r})\mathrm{d}^3r\,. \tag{129}$$

The inversion is obviously

$$f(\mathbf{r}) = \frac{1}{(2\pi)^3}\int \exp(-\mathrm{i}\mathbf{Q}\cdot\mathbf{r})F(\mathbf{Q})\mathrm{d}^3r \tag{130}$$

with one $1/2\pi$ pre-factor for each of the individual coordinates' transform.
An important special case of the 3D FT is that of isotropic functions, where the functions only depend on the absolute values $r = \|\mathbf{r}\| = \sqrt{x^2+y^2+z^2}$ and $Q = \|\mathbf{Q}\|$. In spherical polar coordinates (129) reads:

$$F(Q) = \int_0^{\infty}\mathrm{d}r\int_0^{\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\; r^2\sin\theta\exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r})f(\mathbf{r}) =$$

Because of the symmetry one can assume $\mathbf{Q} = (0,0,Q)$ without loss of generality. Then $\mathbf{Q}\cdot\mathbf{r} = Qr\cos\theta$. In addition, because $f(\mathbf{r})$ does not depend on ϕ the integral over this angle can be carried out: $\int_0^{2\pi}\mathrm{d}\phi = 2\pi$. With this we get:

$$= \int_0^{\infty}\mathrm{d}r\int_0^{\pi}\mathrm{d}\theta\; 2\pi r^2\sin\theta\exp(\mathrm{i}Qr\cos\theta)f(r) =$$

Here, we substitute $\cos\theta \to t$ with the consequence that $\mathrm{d}t = -\sin\theta\mathrm{d}\theta$:

$$= -\int_0^{\infty}\mathrm{d}r\; 2\pi r^2\int_1^{-1}\mathrm{d}t\exp(\mathrm{i}Qrt)f(r) =$$

Because only $\exp(\mathrm{i}Qrt)$ depends on t the integration can be carried out resulting in

$$= \int_0^\infty \frac{\sin(Qr)}{Qr} f(r) 4\pi r^2 \mathrm{d}r = \frac{4\pi}{Q} \int_0^\infty f(r) r \sin(Qr) \mathrm{d}r \,. \tag{131}$$

Similarly it can be shown that the inverse FT is

$$f(r) = \frac{1}{2\pi^2 r} \int_0^\infty F(Q) Q \sin(Qr) \mathrm{d}Q \,. \tag{132}$$

Often the function to be transformed, $f(x)$, corresponds to a physical observable and thus is a real function $f(x) \in \mathbb{R}$. It can be easily shown from the definition (125) that in this case

$$\mathcal{F}_{X|\,x}[f(-x)] = \mathcal{F}_{-X|\,x}[f(x)] = \left(\mathcal{F}_{X|\,x}[f(x)]\right)^* \tag{133}$$

where the star denotes the complex conjugate defined as $(a + b\mathrm{i})^* = a - b\mathrm{i}$. In words: The Fourier transform of the mirror image of a *real* function is the complex conjugate of the original function. This implies that the Fourier transform of an even function is real and that of a real function is even.
In order to describe point scatterers (e.g. nuclei being orders of magnitude smaller than the scattered wave) it is convenient to introduce the delta 'function' as the following limit:

$$\delta(x) = \lim_{a \to 0} \begin{cases} 1/a & \text{for } |x| < a/2 \\ 0 & \text{everywhere else} \end{cases} \,. \tag{134}$$

When the limit is carried out the delta function is everywhere zero except for $x = 0$ where it has an infinite value. What is important is that the area under the function is always one during the whole limiting process. Therefore, it can be concluded that $\int_{-\infty}^{\infty} \delta(x) \mathrm{d}x = 1$. This integral condition and the fact that only the value at $x = 0$ does not vanish are the only characteristics of the delta function. Therefore, different definitions by limits are possible, e.g. by narrowing Gaussian functions. It is somehow justified to say that the delta function is the Black Hole of mathematics, where every individuality of the function is lost as that of a star when it collapses. The property making the delta function interesting is that it is able to pull out a single value of a function from a definite integral:

$$\int_{-\infty}^{\infty} \delta(x) f(x) \mathrm{d}x = \lim_{a \to 0} \frac{1}{a} \int_{-a/2}^{a/2} f(x) \mathrm{d}x =$$

Because in the limit of small a the function does not vary much over the interval $[-a/2, a/2]$ it can be replaced by its value at the centre, $f(0)$:

$$= \lim_{a \to 0} \frac{1}{a} a f(0) = f(0) \,. \tag{135}$$

From the basic property of the delta function (135) follows that the FT as the integral (125) is

$$\mathcal{F}_{X|\,x}[\delta(x)] = \int_{-\infty}^{\infty} \exp(\mathrm{i}Xx) \delta(x) \mathrm{d}x = 1 \,, \tag{136}$$

i.e. the FT of the delta function is the constant function $F(X) = 1$. By writing down the inversion formula (127) one obtains that the iFT of a constant is a delta function:

$$\mathcal{F}^{-1}_{x|\,X}[1] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\mathrm{i}Xx) \mathrm{d}x = \delta(x) \,. \tag{137}$$

(At this point it is important not to forget the 2π !) Because in the last two formulae $\mathrm{i}Xx$ can be replaced by $-\mathrm{i}Xx$ it is of course also true that the iFT of the delta function is a constant $(1/2\pi)$ and the FT of a constant a delta function multiplied by 2π:

$$\mathcal{F}^{-1}_{x|\,X}[\delta(X)] = \frac{1}{2\pi}\,, \tag{138}$$

$$\mathcal{F}_{X|\,x}[1] = 2\pi\delta(x)\,. \tag{139}$$

There are a couple of theorems on definite integrals including FTs which in physics are usually called 'sum rules'. The simplest of these is that for the infinite integral of the FT itself:

$$\begin{aligned}\int_{-\infty}^{\infty}\mathcal{F}_{X|\,x}[f(x)]\mathrm{d}X &= \int_{-\infty}^{\infty}\mathrm{d}X\int_{-\infty}^{\infty}\mathrm{d}x\exp(iXx)f(x)\\ &= \int_{-\infty}^{\infty}f(x)2\pi\delta(x)\mathrm{d}x = 2\pi f(0)\,.\end{aligned} \tag{140}$$

(using that $\int_{-\infty}^{\infty}\mathrm{d}x\exp(iXx)$ is the FT of constant 1, see (139).) Here, it is assumed that the integrals over X and x are interchangeable which may cause problems for certain functions. Often it may be necessary to interpret the integral over X as Cauchy principal value and handle discontinuities in $f(x)$ by imposing Dirichlet's condition, $f(a) = (\lim_{x\searrow a} f(x) + \lim_{x\nearrow a} f(x))/2$. The complementary sum rule is even simpler to derive:

$$\mathcal{F}_{0|\,x}[f(x)] = \int_{-\infty}^{\infty}f(x)\exp(\mathrm{i}0x)\mathrm{d}x = \int_{-\infty}^{\infty}f(x)\mathrm{d}x\,. \tag{141}$$

We see that the infinite integral of a FT corresponds to the value of the original function at zero (up to a factor 2π) and vice versa.
The *convolution*[10] of two functions is defined as follows:

$$f\otimes g(x) = \int_{-\infty}^{\infty}f(t)g(x-t)\mathrm{d}t\,. \tag{142}$$

One of the most important theorems on FTs is that the FT of a convolution is the product of the individual FTs:

$$\mathcal{F}_{X|\,x}[f\otimes g(x)] = \mathcal{F}_{X|\,x}[f(x)]\cdot\mathcal{F}_{X|\,x}[g(x)] = F(X)G(X) \tag{143}$$

or vice versa:

$$\mathcal{F}_{X|\,x}[f(x)g(x)] = \mathcal{F}_{X|\,x}[f(x)]\otimes\mathcal{F}_{X|\,x}[g(x)] = F(X)\otimes G(X)\,. \tag{144}$$

Proof of relation (143): From the definition (125) follows straightforwardly:

$$\begin{aligned}\mathcal{F}_{X|\,x}[f\otimes g(x)] &= \int_{-\infty}^{\infty}\exp(\mathrm{i}Xx)\left(\int_{-\infty}^{\infty}f(t)g(x-t)\mathrm{d}t\right)\mathrm{d}x\\ &= \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}t\exp(\mathrm{i}Xx)f(t)g(x-t)\\ &= \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}t\exp(\mathrm{i}X(x-t))\exp(\mathrm{i}Xt)f(t)g(x-t)\\ &= \int_{-\infty}^{\infty}\mathrm{d}t\int_{-\infty}^{\infty}\mathrm{d}x\exp(\mathrm{i}X(x-t))\exp(\mathrm{i}Xt)f(t)g(x-t)\end{aligned} \tag{145}$$

[10] German: Faltung (sometimes also used in English)

Substituting $x - t \to x'$ leaves the infinite bounds of the first integral unchanged:

$$\begin{aligned} &= \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}x' \exp(\mathrm{i}Xx') \exp(\mathrm{i}Xt) f(t) g(x') \\ &= \int_{-\infty}^{\infty} \exp(\mathrm{i}Xt) f(t) \mathrm{d}t \int_{-\infty}^{\infty} \exp(\mathrm{i}Xx') g(x') \mathrm{d}x' \\ &= \mathcal{F}_{X|x}[f(x)] \cdot \mathcal{F}_{X|x}[g(x)] \end{aligned} \tag{146}$$

Note that the critical step of this proof is the exchange of integrations at line (145). For some 'crazy' functions this may not be allowed and in consequence the convolution theorem does not hold. Nevertheless, for 'physical' functions this is usually no issue.
Closely related to the convolution is the correlator (or correlation function) of two functions:

$$\langle f(0) g(x) \rangle = \int_{-\infty}^{\infty} f(t) g(x+t) \mathrm{d}t \,. \tag{147}$$

By a substitution $t \to -t'$ it is easy to show that

$$\langle f(0) g(x) \rangle = f(-x) \otimes g(x) \tag{148}$$

so that the correlator is just the convolution with one of the functions mirrored. Using the convolution theorem and (133) in succession, it follows that the FT of the correlator of real functions is the product of the FTs of the correlated functions with one FT conjugated, i.e.

$$\mathcal{F}_{X|x}[\langle f(0) g(x) \rangle] = \mathcal{F}_{-X|x}[f(x)] \mathcal{F}_{X|x}[g(x)] = F^*(X) G(X) \,. \tag{149}$$

In physics the most important special case is the autocorrelation function where $f = g$:

$$\mathcal{F}_{X|x}[\langle f(0) f(x) \rangle] = F^*(X) F(X) = |F(X)|^2 \,. \tag{150}$$

This relation, which expresses that the autocorrelation function $\langle f(0) f(t) \rangle$ and the power spectrum $|F(\omega)|^2$ of a signal $f(t)$ are related by a FT, is known as the Wiener-Khintchine theorem.
Only two concrete examples of FTs will be presented here, the exponential/Lorentzian FT pair and the Gaussian function. Interestingly, together with clever use of the rules on Fourier transforms these may cover 90 % of all physical problems. If there is really the necessity to obtain the FT of other functions they may be found in table books as ref. 25. In some cases it may even be more effective to look up the Fourier integral (125) in a table of definite integrals. The (according to the experience of the author) most extensive compilation of such integrals can be found in ref. 26.
Exponential decay: Because the exponential $\exp(-ax)$ diverges for x going to negative infinity one cannot FT it with the two-sided transform. One has either to use a one sided transform or define a function

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \exp(-ax) & \text{for } x > 0 \end{cases} \tag{151}$$

which is cut-off at $x = 0$. This makes also sense in most of the physical contexts, e.g. thinking of the current of a capacitor which is discharged by closing a circuit with a resistor at time zero. The Fourier integral (125) becomes then

$$\int_0^{\infty} \exp(\mathrm{i}Xx) \exp(-ax) \mathrm{d}x = \int_0^{\infty} \exp((\mathrm{i}X - a)x) \mathrm{d}x =$$

The indefinite integral of the exponential is known, $\int \exp(Ax) = \exp(Ax)/A$, thus one continues the calculation

$$= \left.\frac{\exp((\mathrm{i}X - a)x)}{\mathrm{i}X - a}\right|_0^\infty = \frac{1}{a - \mathrm{i}X} = \frac{a}{a^2 + X^2} + \frac{X}{a^2 + X^2}\,\mathrm{i}\,.$$

The real part (which is usually the physically relevant) is a so-called Lorentzian function which peaks at $X = 0$ and has a full width at half maximum of $2a$. Reverting to the interpretation as a decay in time, a is related to the time constant by $a = 1/\tau$. This means that the decay time and the width of the Lorentzian in frequency are inversely proportional.
Gaussian: The bell-shaped curve of the Gaussian (or 'normal') distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{152}$$

decays to both sides rapidly enough and can thus be FTed two-sidedly. The Fourier integral is:

$$\int_{-\infty}^{\infty}\exp(\mathrm{i}Xx)\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right)\mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left(\mathrm{i}Xx - \frac{x^2}{2\sigma^2}\right)\mathrm{d}x =$$

In order to simplify the integral we are going to 'complete the square'[11] inside the exponential by adding $\sigma^2X^2/2$ and compensating this by a factor in front of the integral:

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\sigma^2X^2}{2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2} + \mathrm{i}Xx + \frac{\sigma^2X^2}{2}\right)\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\sigma^2X^2}{2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{(x - \mathrm{i}\sigma^2X)^2}{2\sigma^2}\right)\mathrm{d}x = \end{aligned}$$

The definite integral of the Gaussian distribution always fulfills the normalisation property $(1/\sqrt{2\pi}\sigma)\int_{-\infty}^{\infty}\exp(-(x - A)^2/2\sigma^2)\mathrm{d}x = 1$, irrespectively of its centre A. Therefore, the simple result is:

$$= \exp\left(-\frac{\sigma^2X^2}{2}\right)\,. \tag{153}$$

This means that the FT of a Gaussian is a Gaussian, a property which is called 'self-reciprocity'. It is shared by all Hermite functions of which the Gaussian is the simplest [23]. Nevertheless, the Fourier transform is not the *same* Gaussian but one with the reciprocal standard deviation $1/\sigma$ of the original Gaussian. Also the FT is not normalised to one but its maximum is fixed to one. This in turn causes that the area under the FT Gaussian is $\int_{-\infty}^{\infty}\exp(-\sigma^2X^2/2)\mathrm{d}x = \sqrt{2\pi}/\sigma$.

[11] German: quadratische Ergänzung

References

[1] G. E. Bacon, *Neutron Diffraction* (Clarendon Press, Oxford, 1975).

[2] G. L. Squires, *Introduction to the theory of thermal neutron scattering* (Cambridge University Press, Cambridge, 1978).

[3] S. W. Lovesey: "Theory of Neutron Scattering from Condensed Matter" (Clarendon Press, Oxford, 1984).

[4] W. Brown, *Light Scattering: Principles and Development* (Oxford University Press, Oxford, 1996)

[5] B. Chu, *Laser Light Scattering* (Academic Press, New York, 1974)

[6] H. Z. Cummins and E. R. Pike (eds.), *Photon Correlation Spectroscopy and Velocimetry* (Plenum Press, New York, 1977)

[7] B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley, New York, 1976)

[8] R. Pecora, *Dynamic Light Scattering* (Plenum Press, New York, 1985)

[9] R. Hosemann, A. N. Bagchi: "Direct analysis of diffraction by matter" (North-Holland, Amsterdam, 1962).

[10] L. V. Azaroff, R. Kaplow, N. Kato, R. J. Weiss, A. J. C. Wilson, R. A. Young: "X-ray diffraction" (McGraw-Hill,1974).

[11] O. Kratky, O. Glatter (eds.), *Small Angle X-Ray Scattering* (Academic Press, London, 1982)

[12] Th. Hahn (ed.): "International Tables for Crystallography" (Kluver Academiv Publishers, 2006).

[13] R. Klein: "Interacting Colloidal Suspensions" in P. Lindner and Th. Zemb (eds.), *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter* (Elsevier, Amsterdam, 2002), pp. 351–380.

[14] R. Zorn: "Inelastic Scattering" in T. Brückel, G. Heger, D. Richter, G. Roth, R. Zorn (eds.): "Neutron Scattering" (Forschungszentrum Jülich, 2011), chapter 11.

[15] S. K. Sinha, *J. Phys.: Condens. Matter* **13**, 7511 (2001).

[16] P. N. Pusey, H. M. Fijnaut, A. Vrij, *J. Chem. Phys.* **77**, 4270 (1982).

[17] G. Placzek, *Phys. Rev.*, **86** 377 (1952).

[18] J. L. Yarnell, M. J. Katz, R. G. Wenzel, S. H. Koenig, *Phys. Rev. A*, **7** 2130 (1973).

[19] P. Schofield, *Phys. Rev. Lett.* **4**, 239 (1960).

[20] This is a variant of equation 9.6.34 in M. Abramowitz, I. A. Stegun: "Handbook of Mathematical Functions" (Dover, New York, 1972).

[21] R. Zorn: "Fourier Transforms" in T. Brückel, G. Heger, D. Richter, R. Zorn (eds.): "Neutron Scattering" (Forschungszentrum Jülich, 2008, http://hdl.handle.net/2128/37180), chapter I.

[22] K. B. Howell: "Principles of Fourier Analysis" (Chapman & Hall/CRC, Boca Raton, 2001).

[23] H. J. Weaver: "Theory of discrete and continuous Fourier analysis" (John Wiley & Sons, New York, 1988).

[24] J. Kauppinen, J. Partanen: "Fourier Transforms in Spectroscopy" (Wiley-VCH, Berlin, 2001).

[25] F. Oberhettinger: "Tables of Fourier Transforms and Fourier Transforms of Distributions" (Springer, Berlin, 1990).

[26] A. P. Prudnikov, Yu. A. Brychkov, O. I. Marichev: "Integrals and Series" (Gordon and Breach, New York 1986–1992).

B 1 Crystal Structures and Symmetries

G. Roth
Institute of Crystallography
RWTH Aachen University

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

Introduction

The term "crystal" derives from the Greek κρύσταλλος , which was first used as description of ice and later - in a more general meaning - for transparent minerals with regular morphology (regular crystal faces and edges).

Crystalline solids are thermodynamically stable in contrast to amorphous solids and are characterised by a **regular three-dimensional periodic arrangement of atoms** (ions, molecules) in space. This periodic arrangement makes it possible to determine their structure (atomic positions in 3D space) by **diffraction methods**, using the crystal lattice as a three-dimensional diffraction grating.

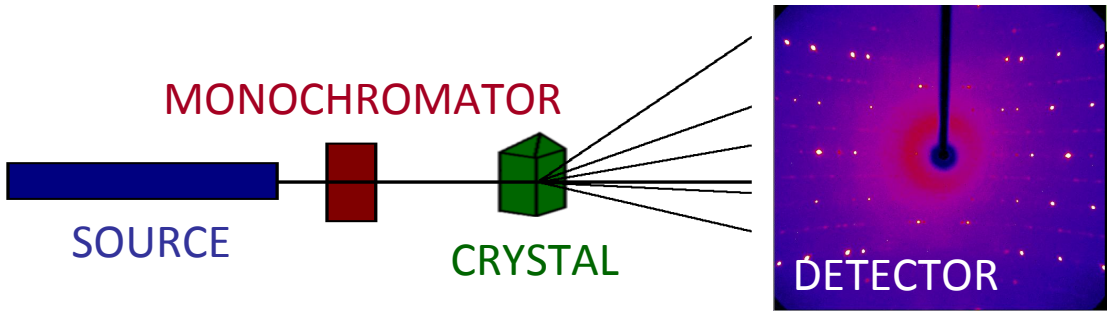


Fig. 1.1: *Sketch of a typical constant wavelength single crystal diffraction experiment. The first such experiment has been conducted by Laue et al. in 1912 (Nobel Prize in Physics 1914).*

The purpose of this chapter is to give a brief introduction into the symmetry concept underlying the description of the crystalline state.

1.1 Crystal lattices

The three-dimensional periodicity of crystals can be represented by the so-called crystal lattice. The repeat unit in form of a parallelepiped - known as the **unit cell** – is defined by 3 non-coplanar basis vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$, whose directions form the reference axes of the corresponding right-handed crystallographic coordinate system. The 6 **lattice parameters** are given as the lengths of the basis vectors $a = |\boldsymbol{a}_1|$, $b = |\boldsymbol{a}_2|$, $c = |\boldsymbol{a}_3|$ and the angles between the basis vectors: angle $(\boldsymbol{a}_1,\boldsymbol{a}_2) = \gamma$, angle $(\boldsymbol{a}_2,\boldsymbol{a}_3) = \alpha$, angle $(\boldsymbol{a}_3,\boldsymbol{a}_1) = \beta$. The faces of the unit cell are named as face $(\boldsymbol{a}_1,\boldsymbol{a}_2)$ = C**,** face $(\boldsymbol{a}_2,\boldsymbol{a}_3)$ = A**,** face $(\boldsymbol{a}_3,\boldsymbol{a}_1)$ = B.

If the vertices of all repeat units (unit cells) are replaced by points, the result is the crystal lattice in the form of a **point lattice**. Each lattice point is given by a vector $\boldsymbol{a} = u\boldsymbol{a}_1+v\boldsymbol{a}_2+w\boldsymbol{a}_3$, with u, v, w being integers. As a symmetry operation of parallel displacement, $\boldsymbol{a}$ – also known as **translation vector** – maps the atomic arrangement of the crystal (crystal structure) onto itself.

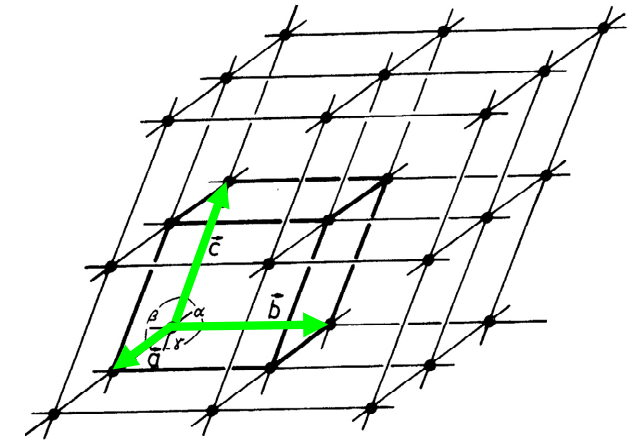


Fig. 1.2: *Notation for a unit cell (basis vectors* $\boldsymbol{a_1}$, $\boldsymbol{a_2}$, $\boldsymbol{a_3}$, *or a, b, c) and a point lattice.*

A **lattice point** is labelled "*uvw*", according to the coefficients (integers) of the translation vector

$$\boldsymbol{a} = u\,\boldsymbol{a_1} + v\,\boldsymbol{a_2} + w\,\boldsymbol{a_3} \qquad 1.1$$

from the origin to the lattice point. A **lattice direction** - given by the symbol [*uvw*] - is defined by the direction of the corresponding translation vector.

A plane passing through three lattice points is known as a **lattice plane**. Since all lattice points are equivalent (by translation symmetry) there will be infinitely many parallel planes passing through all the other points of the lattice. Such a set of equally spaced planes is known as a **set of lattice planes**. If the first plane from the origin of a set of lattice planes makes intercepts a_1/h, a_2/k, a_3/l on the axes, where h, k, l are integers, then the **Miller indices** of this set of lattice planes are (hkl), the three coefficients h, k, l are conventionally enclosed in parentheses.

The equation of lattice planes can be written in intercept form as

$$(h\mathrm{x}/a_1) + (k\mathrm{y}/a_2) + (l\mathrm{z}/a_3) = \mathrm{n}, \qquad 1.2$$

where n is an integer. If n = 0 the lattice plane passes through the origin; if n = 1 the plane makes intercepts a_1/h, a_2/k, a_3/l on the axes; if n = 2 the intercepts are $2a_1/h$, $2a_2/k$, $2a_3/l$; and so on.

Complementary to the crystal lattice, the so-called reciprocal lattice may be constructed, which is a useful tool for understanding the geometry of diffraction experiments. The reciprocal lattice can be thought of as the result of diffraction (of X-rays, neutrons, electrons etc.) from the crystal lattice ('direct lattice'). The points on the diffraction pattern in Fig. 1.1 (right) are actually points of the reciprocal lattice recorded during the diffraction experiment. Their nodes are indexed by the Miller-indices hkl in the same way as the nodes of the direct lattice are indexed by uvw:

$$\boldsymbol{\tau} = h\,\boldsymbol{\tau_1} + k\,\boldsymbol{\tau_2} + l\,\boldsymbol{\tau_3}. \qquad 1.3$$

The basis vectors $\boldsymbol{\tau}$ of the reciprocal lattice can be calculated from those of the direct cell by:

$$\boldsymbol{\tau}_i = (\mathbf{a}_j \times \mathbf{a}_k)/V_c, \qquad 1.4$$

where × means the cross product, and $V_c = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$ is the volume of the unit cell.

Here is a compilation of some properties of the reciprocal lattice:

- Each reciprocal lattice vector is perpendicular to two real space vectors: $\boldsymbol{\tau}_i \perp \mathbf{a}_j$ and $\mathbf{a}_k$ (for $i \neq j, k$)
- The lengths of the reciprocal lattice vectors are $|\boldsymbol{\tau}_i| = 1/V_c \cdot |\mathbf{a}_j| \cdot |\mathbf{a}_k| \cdot \sin\angle(\mathbf{a}_j, \mathbf{a}_k)$.
- Each point *hkl* in the reciprocal lattice refers to a set of planes (*hkl*) in real space.
- The direction of the reciprocal lattice vector $\boldsymbol{\tau}$ is normal to the (*hkl*) planes and its length is reciprocal to the interplanar spacing d_{hkl}: $|\boldsymbol{\tau}| = 1/d_{hkl}$.
- Duality principle: The reciprocal lattice of the reciprocal lattice is the direct lattice.

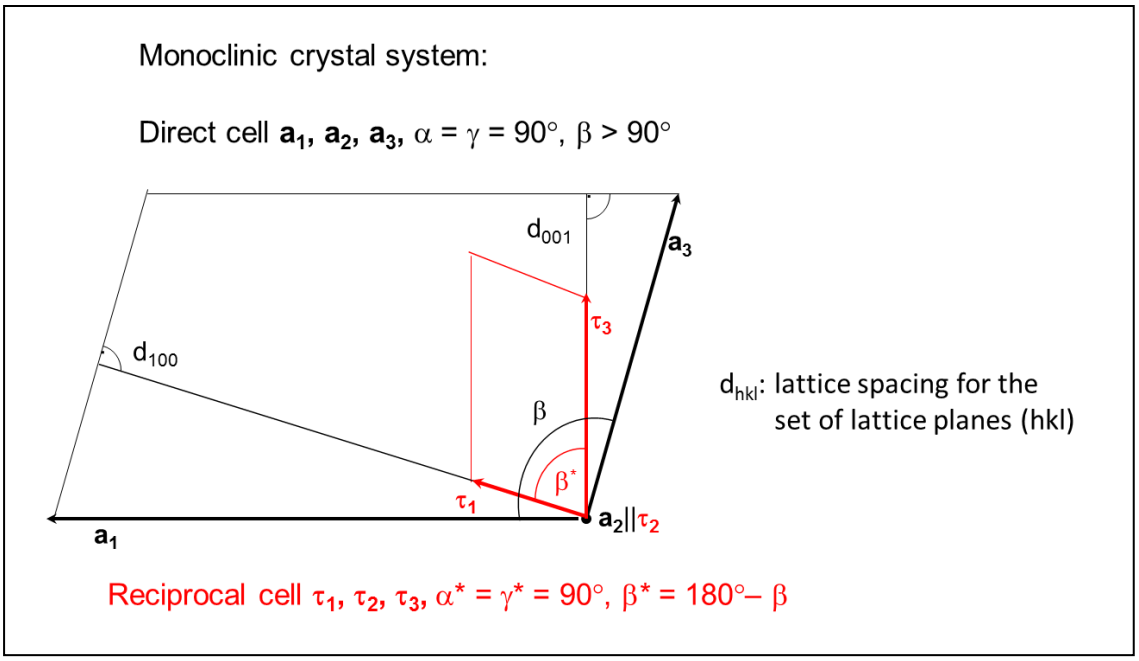


Fig. 1.3: Direct and corresponding reciprocal unit cell.

1.2 Crystallographic coordinate systems

The description of a crystal structure consists first of the choice of a unit cell as the smallest repeat unit of the crystal with its basis vectors. In this way a crystal-specific coordinate system is defined which is used to localize all the atoms in the unit cell. While - in physics and chemistry - Cartesian coordinate systems are frequently used, crystallographers often use non-orthogonal and non-orthonormal coordinate systems.

The conventional crystallographic coordinate systems are based on the symmetry of the crystals. In three dimensions there are **7** different **crystal systems** and hence 7 crystallographic coordinate systems:

Crystal system	Minimum symmetry	Conventional unit cell
triclinic	1 or $\bar{1}$	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$
monoclinic (unique axis b)	one diad – 2 or m ($\parallel Y$)	$a \neq b \neq c$; $\alpha=\gamma=90°$, $\beta>90°$
orthorhombic	three mutually perpendicular diads – 2 or m ($\parallel X$, Y and Z)	$a \neq b \neq c$; $\alpha=\beta=\gamma=90°$
tetragonal	one tetrad – 4 or $\bar{4}$ ($\parallel Z$)	$a = b \neq c$; $\alpha=\beta=\gamma=90°$
trigonal (hexagonal cell)	one triad – 3 or $\bar{3}$ ($\parallel Z$)	$a = b \neq c$; $\alpha=\beta=90°$, $\gamma=120°$
hexagonal	one hexad – 6 or $\bar{6}$ ($\parallel Z$)	$a = b \neq c$; $\alpha=\beta=90°$, $\gamma=120°$
cubic	four triads – 3 or $\bar{3}$ ($\parallel$space diagonals of cube)	$a = b = c$; $\alpha=\beta=\gamma=90°$

The choice of the origin of the coordinate system is free in principle, but for convenience it is usually chosen at a centre of symmetry (inversion centre), if present, otherwise in a point of high symmetry.

In order to complete the symmetry conventions of the coordinate systems it is necessary to add to the 7 so-called primitive unit cells of the crystal systems (primitive lattice types with only one lattice point per unit cell) 7 centred unit cells with two, three or four lattice points per unit cell (centred lattice types). These centred unit cells are consequently two, three or four times larger than the smallest repeat units of the crystals. The resulting **14 Bravais lattice types** with their centering conditions are collected in Fig. 1.4.

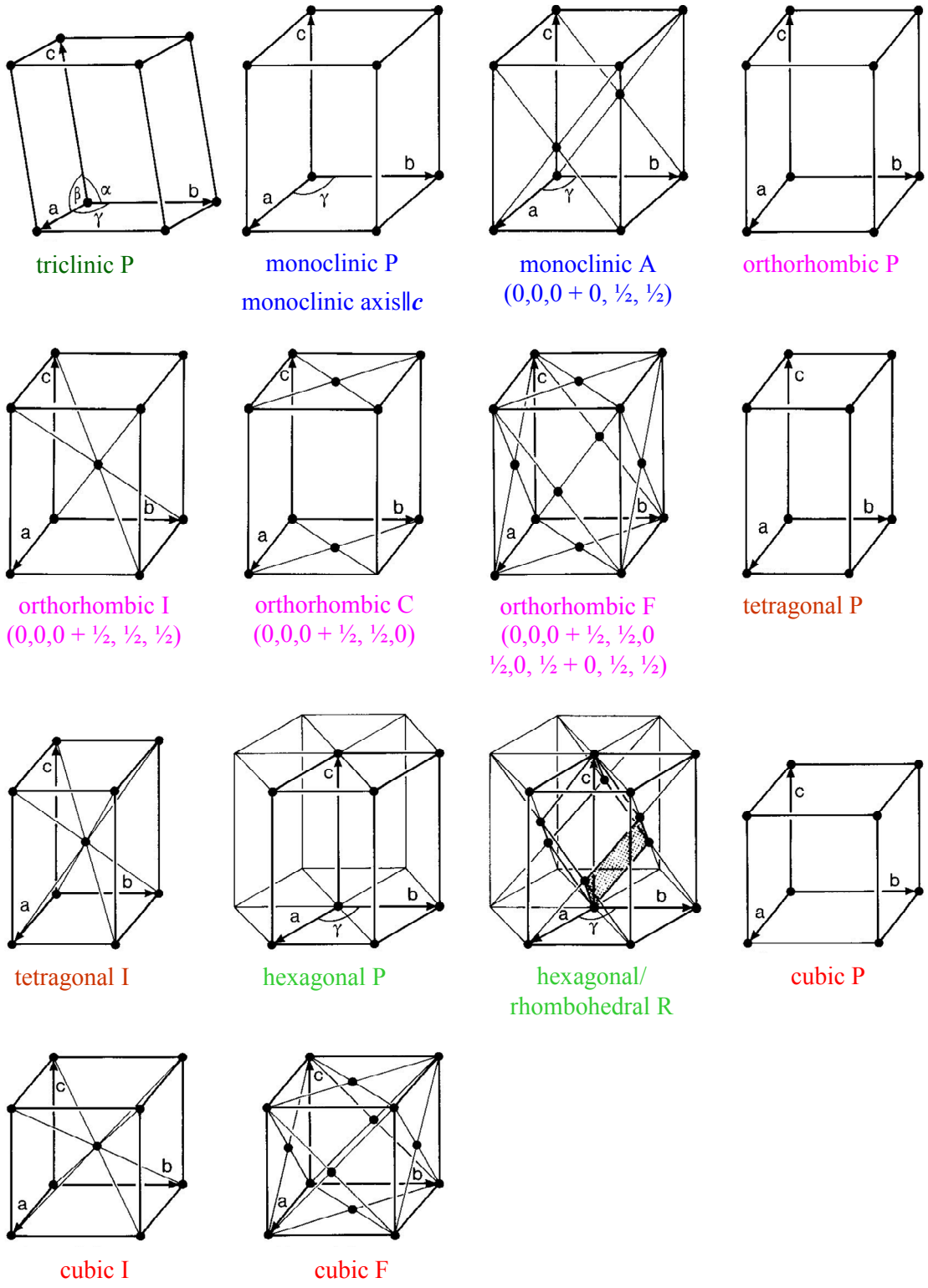


Fig. 1.4: *The 14 Bravais lattices consisting of the 7 primitive lattices P for the 7 crystal systems with only one lattice point per unit cell + the 7 centred (multiple) lattices A, B, C, I, R and F with 2, 3 and 4 lattice points per unit cell.*

1.3 Symmetry-operations and -elements

The **symmetry operations** of a crystal are isometric transformations or motions, i.e. mappings which preserve distances and, hence, also angles and volumes. An object and its transformed object superimpose in a perfect manner, they are indistinguishable.

The simplest crystallographic symmetry operation is the **translation**, which is a parallel displacement of the crystal by a translation vector $\boldsymbol{a}$ (see chapter 1.1). There is no fixed point, the entire lattice is shifted and therefore, theoretically, the crystal lattice is considered to be infinite.

Crystallographic **rotations** n around an axis by an angle $\varphi = 360°/n$ (n-fold rotations) and **rotoinversions** (combination of rotations and inversions) $\bar{n}$ are called point symmetry operations because they leave at least one point of space invariant (at least one fixed point). An important fact of crystallographic symmetry is the restriction of the rotation angles by the three-dimensional crystal lattice to $\varphi = 360°$ (n = 1), 180° (n = 2), 120° (n = 3), 90° (n = 4), 60° (n = 6). Only for these crystallographic rotations the space can be covered completely without gaps and overlaps. The rotoinversion $\bar{n} = \bar{1}$ is an **inversion** in a point, $\bar{n} = \bar{2} \equiv m$ (mirror) describes a **reflection** across a plane.

The combination of n-fold rotations with $(m/n)\cdot\boldsymbol{a}$ translation components (m < n) ∥ to the rotation axis leads to the so-called **screw rotations** n_m, e.g. 2_1, 3_2, 4_2, 6_5. These symmetry operations have no fixed points.

The combination of a reflection through a plane (glide plane) with translation components (glide vectors) of $\boldsymbol{a_1}/2$, $\boldsymbol{a_2}/2$, $\boldsymbol{a_3}/2$, $(\boldsymbol{a_1}+\boldsymbol{a_2})/2$, … ∥ to this plane are known as **glide reflections** a, b, c, n, …, d. Again no fixed points exist for these symmetry operations.

In addition to the symmetry operations which represent isometric motions of an object, symmetry can also be described in (static) geometrical terms by **symmetry elements**. They form the geometrical locus, oriented in space, on which a symmetry operation is performed (line for a rotation, plane for a reflection, and point for an inversion) together with a description of this operation. Symmetry elements are mirror planes, glide planes, rotation axes, screw axes, rotoinversion axes and inversion centres. The geometrical descriptions of the crystallographic symmetry operations are illustrated in Figs. 1.5-1.7.

A symmetry operation transforms a point X with coordinates x, y, z (according to a position vector $\boldsymbol{X} = x\boldsymbol{a_1} + y\boldsymbol{a_2} + z\boldsymbol{a_3}$) into a symmetrically equivalent point X' with coordinates x', y', z' mathematically by the linear equations

$$\begin{aligned} x' &= W_{11}x + W_{12}y + W_{13}z + w_1 \\ y' &= W_{21}x + W_{22}y + W_{23}z + w_2 \\ z' &= W_{31}x + W_{32}y + W_{33}z + w_3 \end{aligned} \qquad 1.5$$

with w_1, w_2, w_3 constituting the translational part of the symmetry operation.

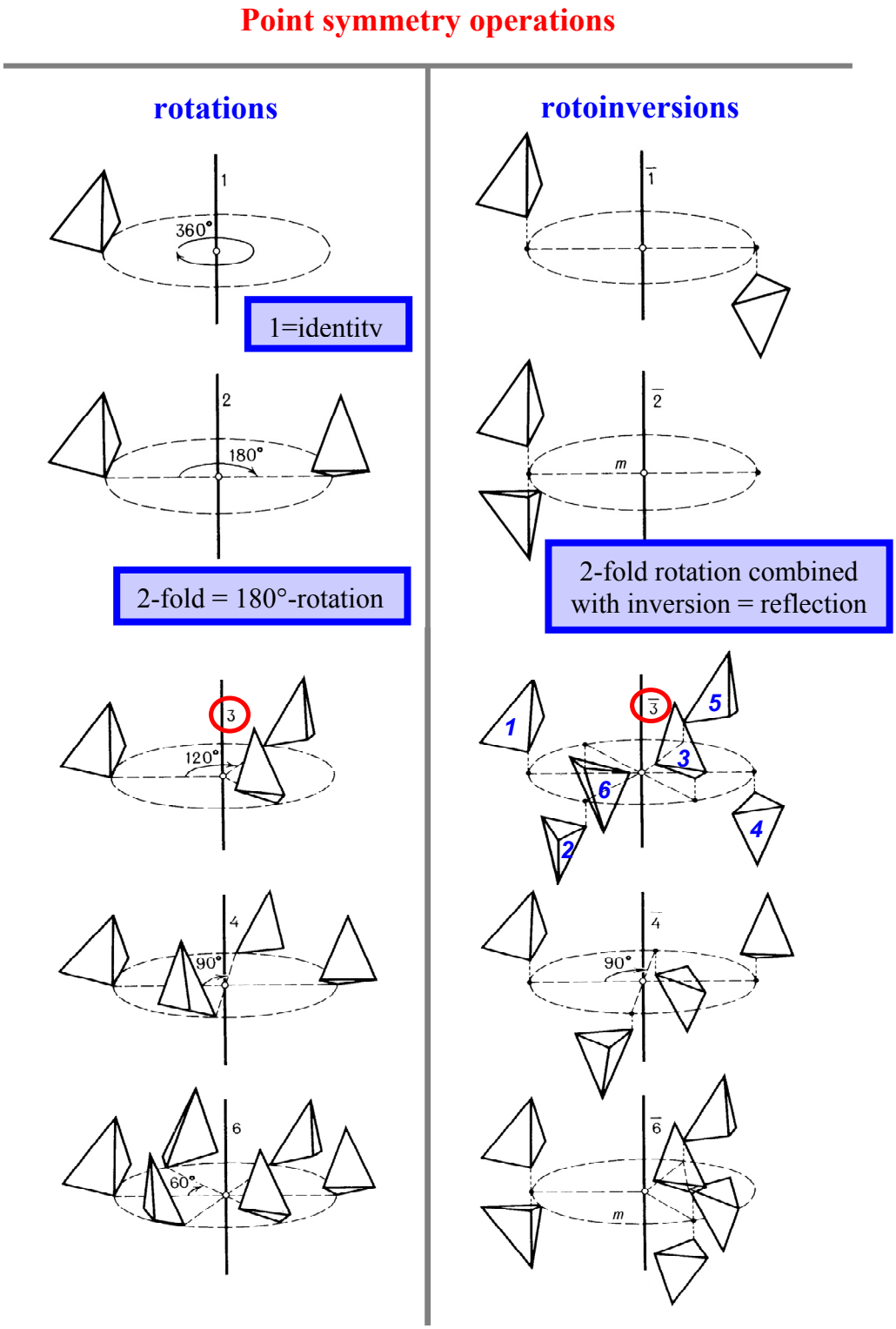


Fig. 1.5: *Rotations: n=1 (identity), n=2 (rot. angle 180°), n=3 (120°), n=4 (90°), n=6 (60°). Rotoinversions: $\bar{1}$ (inversion), $\bar{2} \equiv m$ (reflection), 3=3+1, $\bar{4}$, $\bar{6}$ = 3/m.*

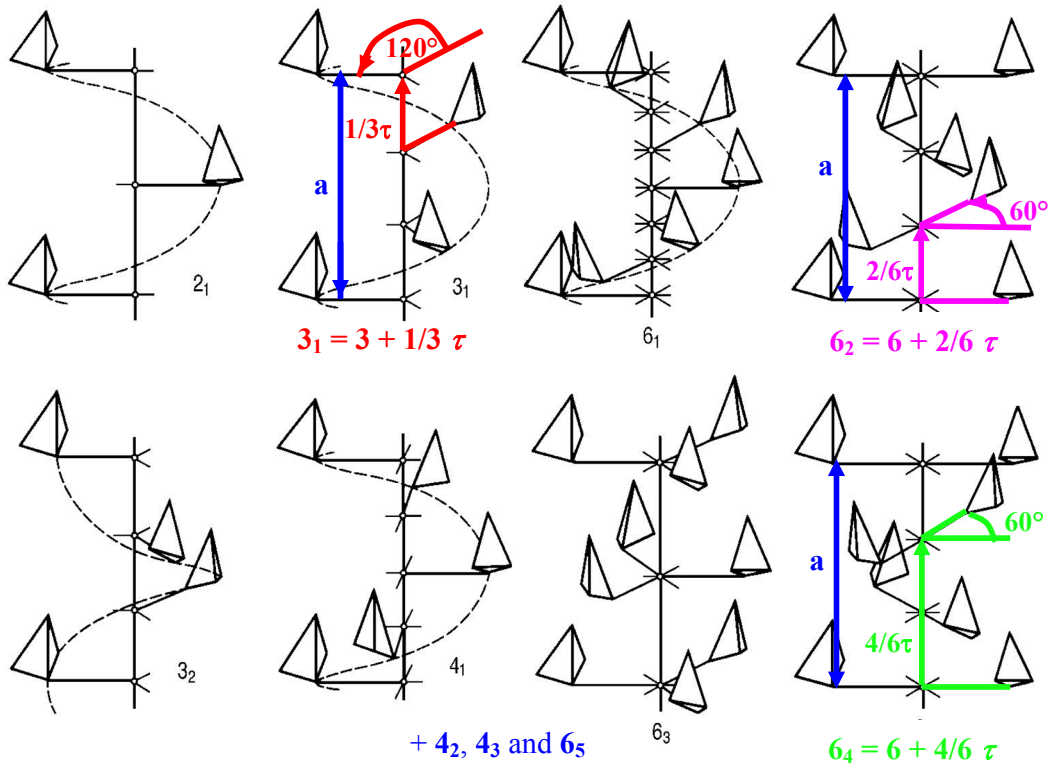


Fig. 1.6: *Screw rotations n_m: combination of rotations n and translation components (m/n)·a ∥ to the rotation axis.*

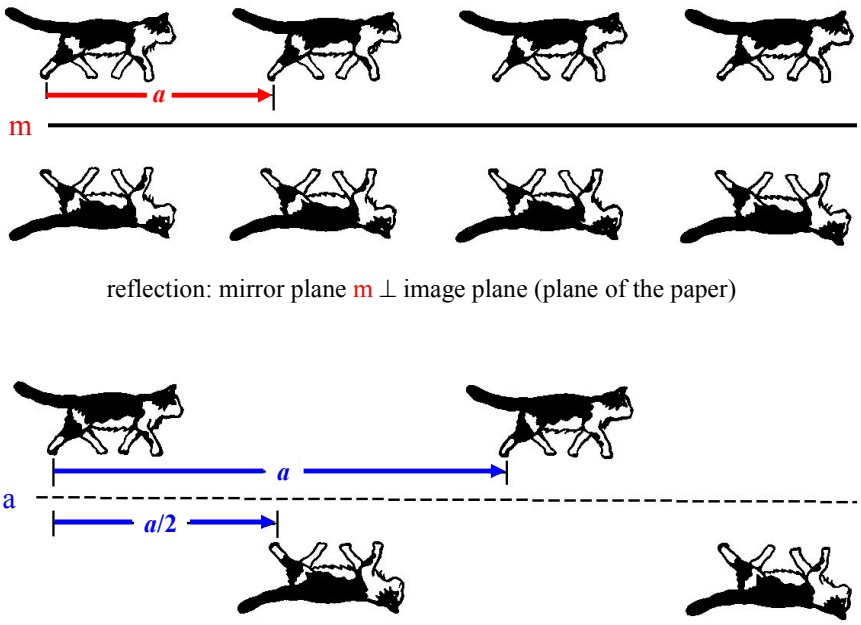


Fig. 1.7: *Examples of reflections and glide reflections.*

The above equation, re-written in matrix notation:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix} \circ \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}; \quad \boldsymbol{X}' = \mathbf{W} \circ \boldsymbol{X} + \boldsymbol{w} = (\mathbf{W}, \mathbf{w}) \circ \boldsymbol{X} \qquad 1.6$$

The (3×3) matrix **W** is the rotational part and the (3×1) column matrix ***w*** the translational part of the symmetry operation. The two parts **W** and ***w*** can be assembled into an augmented (4×4) matrix ***W*** according to

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & W_{13} & w_1 \\ W_{21} & W_{22} & W_{23} & w_2 \\ W_{31} & W_{32} & W_{33} & w_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \circ \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \boldsymbol{W} \circ \boldsymbol{X} \qquad 1.7$$

Since every symmetry transformation is a "rigid-body" motion, the determinant of all matrices **W** and ***W*** is det **W** = det ***W*** = ± 1 (+ 1: preservation of handedness; - 1: change of handedness of the object).

The sequence of two symmetry operations (successive application) is given by the product of their matrices $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$:

$$\boldsymbol{W_3} = \boldsymbol{W_1} \circ \boldsymbol{W_2} \qquad 1.8$$

where $\boldsymbol{W_3}$ is again a symmetry operation.

1.4 Crystallographic point groups and space groups

The symmetry of a crystal and of its crystal structure can be described by mathematical group theory. The symmetry operations are the group elements of a crystallographic group G and the combination of group elements is the successive execution of symmetry operations. All possible combinations of crystallographic point-symmetry operations in three-dimensional space lead to exactly 32 crystallographic point groups (≡ crystal classes) which all are of finite order (the maximum order is 48 for the cubic crystal class $m\bar{3}m$). For the different crystal systems they are represented by stereographic projections in Fig. 1.8. There are two types of group symbols in use: For each crystal class the corresponding Schoenflies symbol is given at the bottom left and the Hermann-Mauguin (international) symbol at the bottom right. A maximum of 3 independent main symmetry directions ("Blickrichtungen") is sufficient to describe the complete symmetry of a crystal. These symmetry directions are specifically defined for the 7 crystal systems (Hermann-Mauguin symbols). As an example, the symmetry directions of the cubic system are shown in Fig. 1.9.

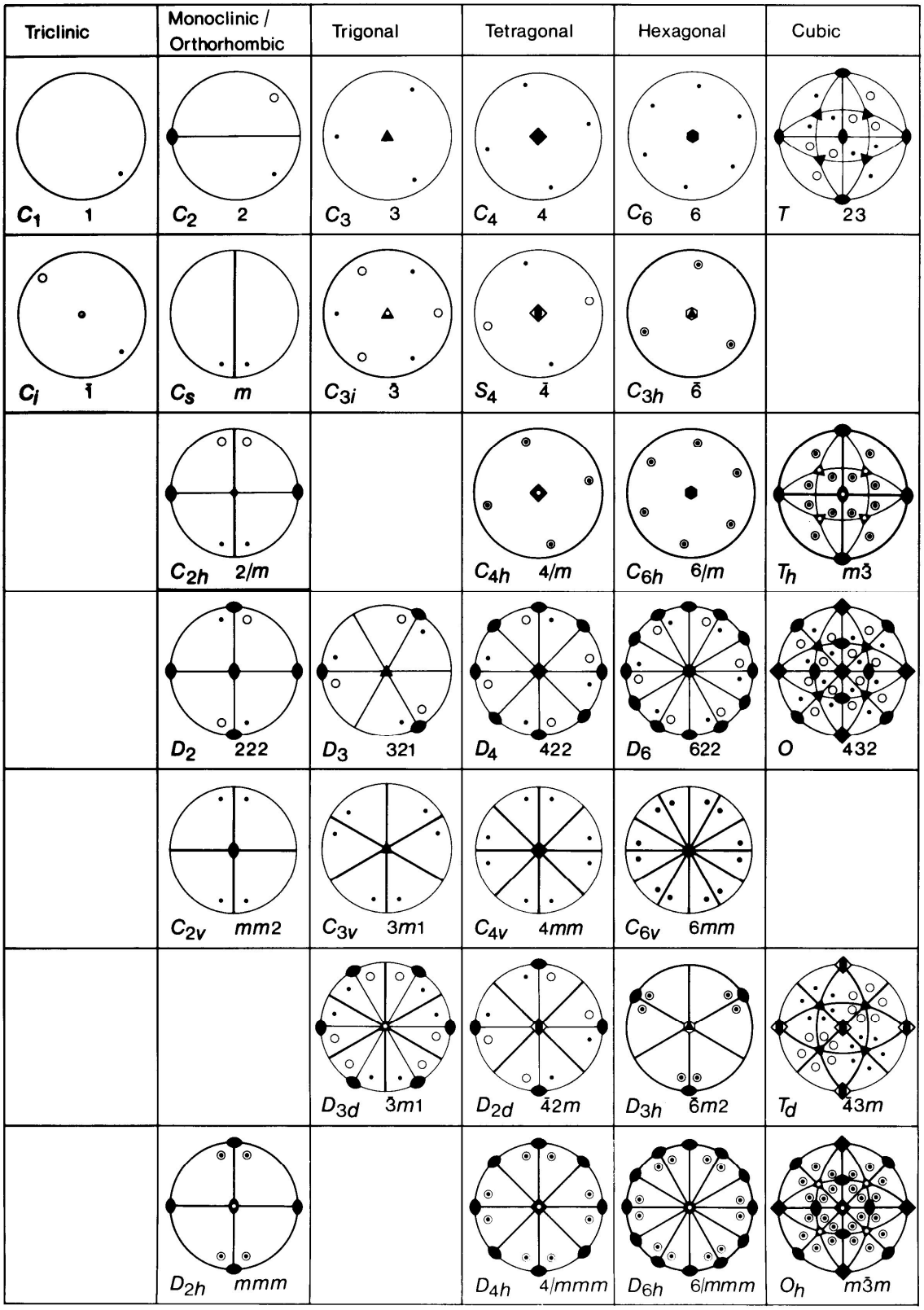


Fig. 1.8: *The 32 crystallographic point groups (crystal classes) in three-dimensional space represented by their stereographic projections. The group symbols are given according to Schoenflies (bottom left) and to Hermann-Mauguin (bottom right).*

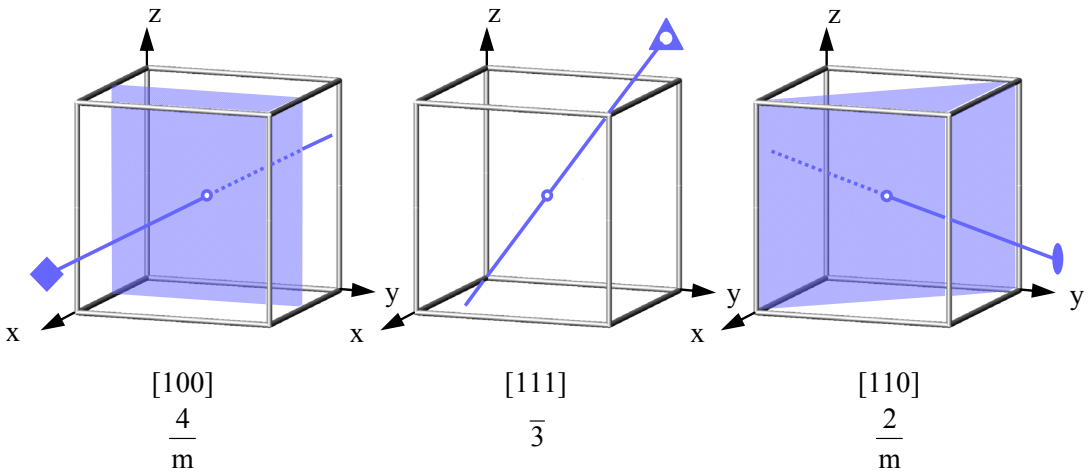


Fig. 1.9: *Symmetry directions ("Blickrichtungen") of the cubic lattice (a=b=c, α=β=γ=90°). Along [100]: 4/m, along [111]: $\bar{3}$, along [110]: 2/m.*

In three dimensions all possible combinations of the point symmetries of the 32 crystallographic point groups with the lattice translations of the 14 Bravais lattices lead to exactly 230 space groups, all of infinite order. As already mentioned, the addition of translations to the point symmetries results in new symmetry operations: Screw rotations and glide reflections. The conventional graphical symbols for the symmetry elements according to the International Tables for Crystallography Vol. A (2002) [1] are shown in Fig. 1.10.

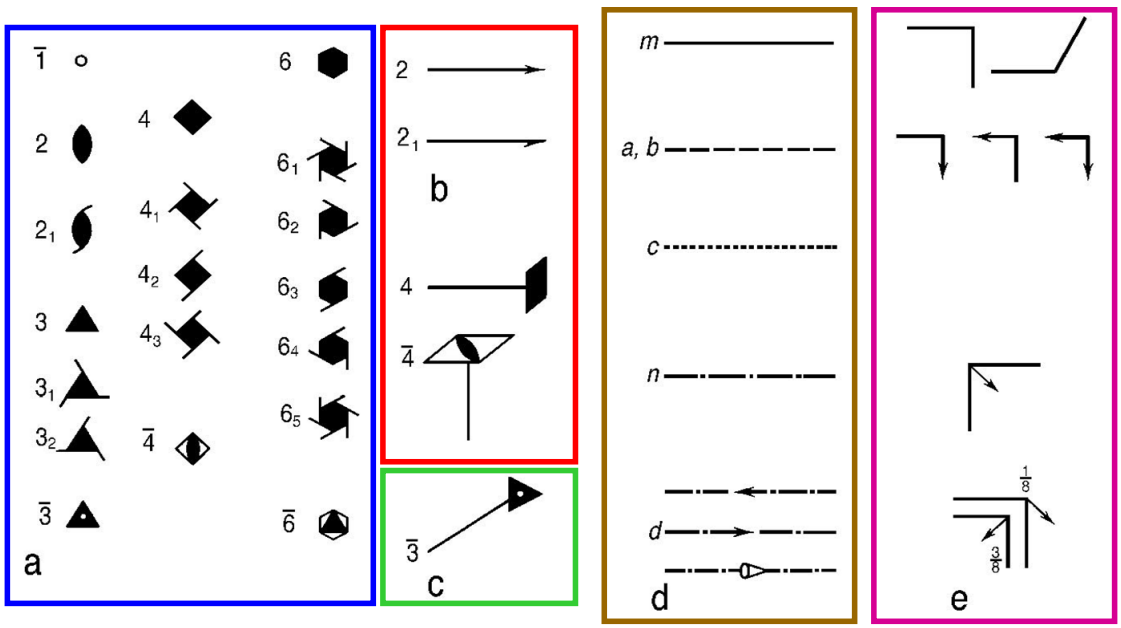


Fig. 1.10: *Conventional graphical symbols for symmetry elements:*
- symmetry axes: (a) perpendicular, (b) parallel, and (c) inclined to the plane;
- symmetry planes: (d) perpendicular and (e) parallel to the image plane.

1.5 Quasicrystals

Since the pioneering work of Shechtman et al [2] published in 1984 and honoured by the 2011 Nobel-Prize in Physics it is accepted that the crystalline state with its 3D periodic arrangement of atoms in a lattice is not the only long-range ordered ground state of matter. This quasi crystalline state also follows strict construction rules and exhibits long range order, but the rules are no longer based on the lattice concept.

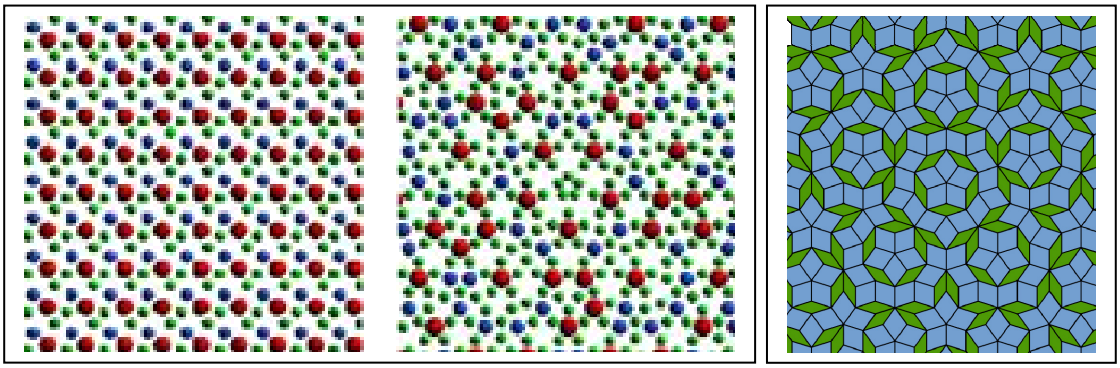


Fig. 1.11: 2D-analogues of a crystalline (left) and a quasi-crystalline structure (center) [3], Penrose tiling of a plane by two different rhombs (right) [4].

The description of quasicrystals is closely related to the so called Penrose-tilings which are a way to cover the plane completely and without overlap by a long range ordered, non-periodic arrangement of (in the case shown in Fig. 1.11) two different geometric shapes (here: rhombs). As a result of the lack of translation symmetry, the "crystallographically forbidden" rotation axes (5-fold, 8-fold, 10-fold etc., more precisely: forbidden as part of a 3D-space group symmetry) may occur in quasi crystals and also show up as symmetries of the outer shape (Fig. 1.12 left) and diffraction patterns of quasi crystals (center).

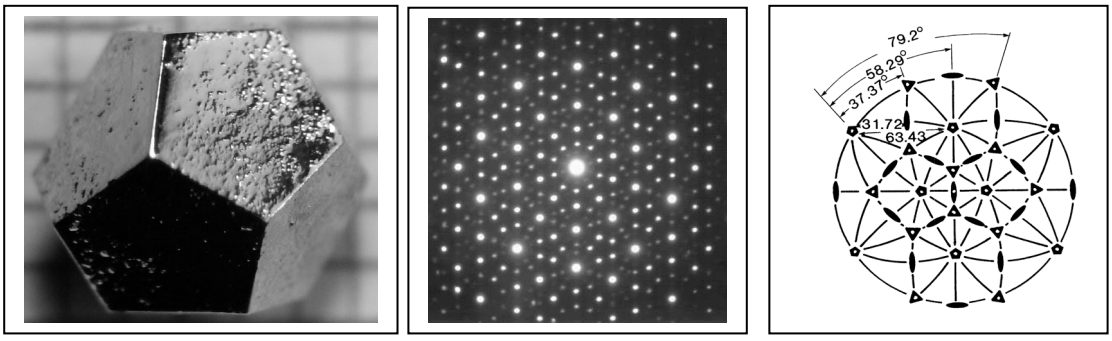


Fig. 1.12: Icosahedral quasi crystal HoMgZn (left) [5], electron diffraction pattern taken along the -5 rotoinversion axis (center) [5] and stereographic projection of the icosahedral point symmetry group m-3-5 [2]

1.6 Application: Structure description of $YBa_2Cu_3O_{7-\delta}$

The crystal structure determination with atomic resolution is achieved by diffraction experiments with X-rays, electron or neutron radiation. As an example, the results of a structure analysis by neutron diffraction on a single crystal of the ceramic high-T_C superconductor $YBa_2Cu_3O_{7-\delta}$ with T_C = 92 K are presented [6]. The atomic arrangement of the orthorhombic structure, space group P m m m, and the temperature-dependent electrical resistivity are shown in Fig. 1.13.

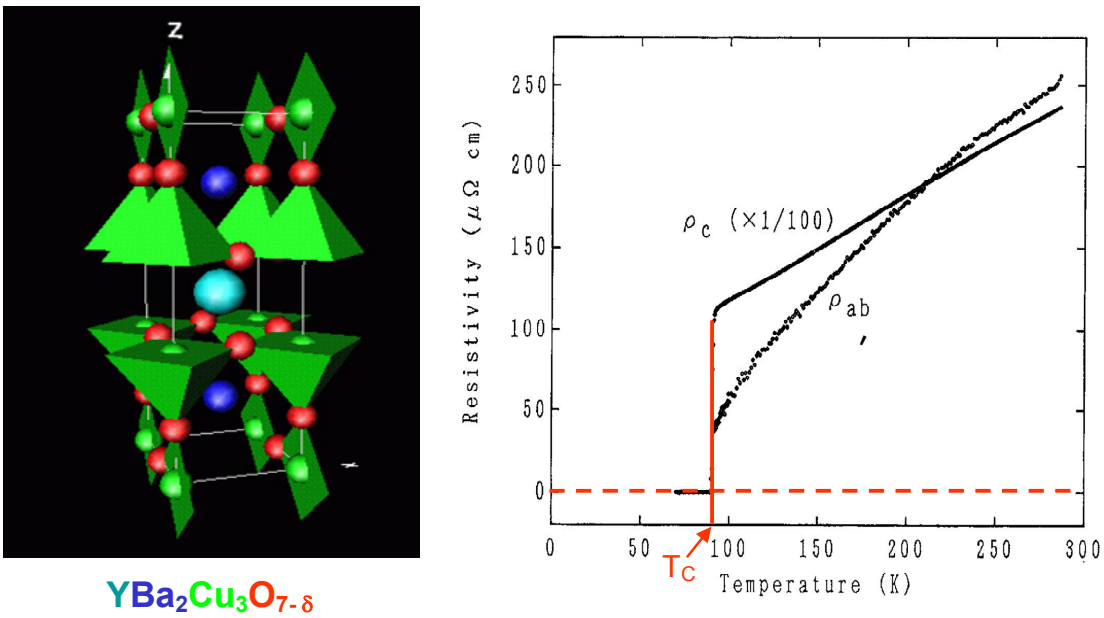


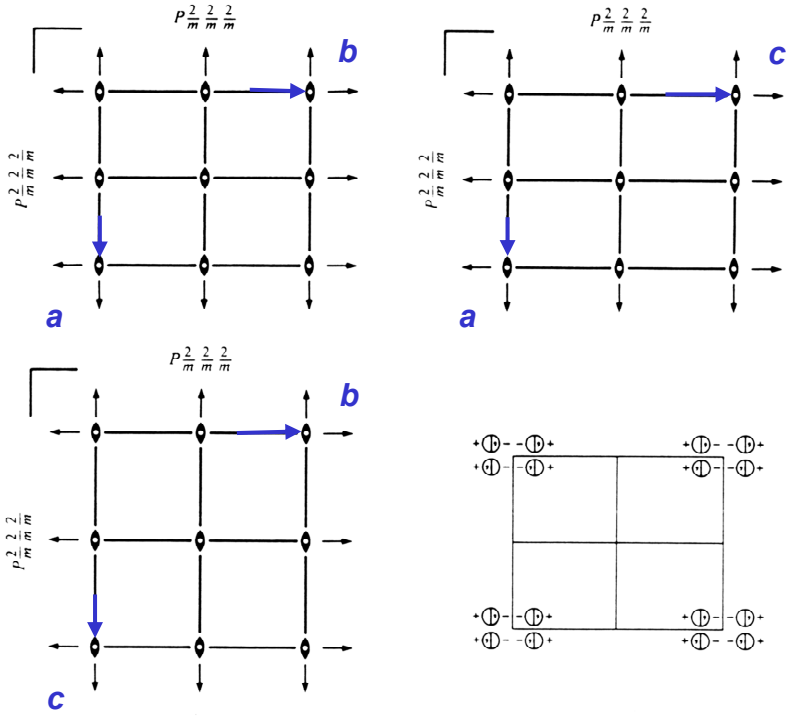
Fig. 1.13: *Crystal structure (unit cell) of $YBa_2Cu_3O_{7-\delta}$ with the CuO_x-polyhedra (left) and the electrical resistivity as a function of temperature ∥ and ⊥ to the [001] direction (right).*

The crystal structure contains two different Cu-O polyhedra (green): CuO_5-tetragonal pyramids and CuO_4-squares. The pyramids share corners in 2D and form double layers, the charge carriers responsible for superconductivity are supposed to be located in these double layers.

Information from the international tables on the relative locations and orientations of the symmetry elements (symmetry operations 1, 2_z, 2_y, 2_x, $\bar{1}$, m_z, m_y, m_x) of the orthorhombic space group P m m m, together with the choice of the origin (in an inversion centre), is shown in Fig. 1.14. The general position (site symmetry 1) of multiplicity 8 and all special positions with their site symmetries are listed in Fig. 1.15. There are no special reflection conditions for this space group.

$P\,m\,m\,m$ **D_{2h}^{1}** ***m m m*** **Orthorhombic**

No. 47 *P 2/m 2/m 2/m* Patterson symmetry ***P m m m***



Origin at centre ($m\,m\,m$)

Asymmetric unit $0 \le x \le \frac{1}{4}$; $0 \le y \le \frac{1}{4}$; $0 \le z \le \frac{1}{4}$

Symmetry operations

(1) 1 (2) 2 $0,0,z$ (3) 2 $0,y,0$ (4) 2 $x,0,0$
(5) $\bar{1}$ $0,0,0$ (6) m $x,y,0$ (7) m $x,0,z$ (8) m $0,y,z$

Fig. 1.14: *Description of the orthorhombic space group P m m m in [1].*

CONTINUED $YBa_2Cu_3O_{7-\delta}$ No. 47 *Pmmm*

Generators selected (1); $t(1,0,0)$; $t(0,1,0)$; $t(0,0,1)$; (2); (3); (5)

Positions

Multiplicity, Wyckoff letter, Site symmetry			Coordinates				Reflection conditions
							General:
8	α	1	(1) x,y,z	(2) $\bar{x},\bar{y},z$	(3) $\bar{x},y,\bar{z}$	(4) $x,\bar{y},\bar{z}$	no conditions
			(5) $\bar{x},\bar{y},\bar{z}$	(6) $x,y,\bar{z}$	(7) $x,\bar{y},z$	(8) $\bar{x},y,z$	
							Special: no extra conditions
4	z	$..m$	$x,y,\frac{1}{2}$	$\bar{x},\bar{y},\frac{1}{2}$	$\bar{x},y,\frac{1}{2}$	$x,\bar{y},\frac{1}{2}$	
4	y	$..m$	$x,y,0$	$\bar{x},\bar{y},0$	$\bar{x},y,0$	$x,\bar{y},0$	
4	x	$.m.$	$x,\frac{1}{2},z$	$\bar{x},\frac{1}{2},z$	$\bar{x},\frac{1}{2},\bar{z}$	$x,\frac{1}{2},\bar{z}$	
4	w	$.m.$	$x,0,z$	$\bar{x},0,z$	$\bar{x},0,\bar{z}$	$x,0,\bar{z}$	
4	v	$m..$	$\frac{1}{2},y,z$	$\frac{1}{2},\bar{y},z$	$\frac{1}{2},y,\bar{z}$	$\frac{1}{2},\bar{y},\bar{z}$	
4	u	$m..$	$0,y,z$	$0,\bar{y},z$	$0,y,\bar{z}$	$0,\bar{y},\bar{z}$	
2	t	$mm2$	$\frac{1}{2},\frac{1}{2},z$	$\frac{1}{2},\frac{1}{2},\bar{z}$			
2	s	$mm2$	$\frac{1}{2},0,z$	$\frac{1}{2},0,\bar{z}$			
2	r	$mm2$	$0,\frac{1}{2},z$	$0,\frac{1}{2},\bar{z}$			
2	q	$mm2$	$0,0,z$	$0,0,\bar{z}$			
2	p	$m2m$	$\frac{1}{2},y,\frac{1}{2}$	$\frac{1}{2},\bar{y},\frac{1}{2}$			
2	o	$m2m$	$\frac{1}{2},y,0$	$\frac{1}{2},\bar{y},0$			
2	n	$m2m$	$0,y,\frac{1}{2}$	$0,\bar{y},\frac{1}{2}$			
2	m	$m2m$	$0,y,0$	$0,\bar{y},0$			
2	l	$2mm$	$x,\frac{1}{2},\frac{1}{2}$	$\bar{x},\frac{1}{2},\frac{1}{2}$			
2	k	$2mm$	$x,\frac{1}{2},0$	$\bar{x},\frac{1}{2},0$			
2	j	$2mm$	$x,0,\frac{1}{2}$	$\bar{x},0,\frac{1}{2}$			
2	i	$2mm$	$x,0,0$	$\bar{x},0,0$			
1	h	mmm	$\frac{1}{2},\frac{1}{2},\frac{1}{2}$				
1	g	mmm	$0,\frac{1}{2},\frac{1}{2}$				
1	f	mmm	$\frac{1}{2},\frac{1}{2},0$				
1	e	mmm	$0,\frac{1}{2},0$				
1	d	mmm	$\frac{1}{2},0,\frac{1}{2}$				
1	c	mmm	$0,0,\frac{1}{2}$				
1	b	mmm	$\frac{1}{2},0,0$				
1	a	mmm	$0,0,0$				

Fig. 1.15: *General and special positions (coordinates of all symmetrically equivalent positions) of space group P m m m with their site symmetries and multiplicities [1]. The special positions occupied by atoms of the $YBa_2Cu_3O_{7-\delta}$ structure are highlighted by frames.*

The atomic parameters of the structure refinement of $YBa_2Cu_3O_{6.96}$ at room temperature [6] are given in the following Table:

Atomic positions of $YBa_2Cu_3O_{6.96}$ orthorhombic, space group type P *2/m 2/m 2/m* *a* = 3.858 Å, *b* = 3.846 Å, *c* = 11.680 Å (at room temperature)					
atom/ion	**multiplicity**	**site symmetry**	**x**	**y**	**z**
$Cu1/Cu^{2+}$	1	*2/m 2/m 2/m*	0	0	0
$Cu2/Cu^{2+}$	2	*m m 2*	0	0	0.35513(4)
Y/Y^{3+}	1	*2/m 2/m 2/m*	½	½	½
Ba/Ba^{2+}	2	*m m 2*	½	½	0.18420(6)
$O1/O^{2-}$	2	*m m 2*	0	0	0.15863(5)
$O2/O^{2-}$	2	*m m 2*	0	½	0.37831(2)
$O3/O^{2-}$	2	*m m 2*	½	0	0.37631(2)
$O4/O^{2-}$	1	*2/m 2/m 2/m*	0	½	0

References

[1] International Tables for Crystallography Vol. A, *Space-group Symmetry*, edited by Th. Hahn, Dordrecht: Kluwer Academic Publishers (5. Edition, 2002)

[2] D. Shechtman, I. Blech, D. Gratias, J. Cahn: "*Metallic Phase with Long-Range Orientational Order and No Translational Symmetry*". Physical Review Letters **53**, 1951 (1984)

[3] S. Curtarolo, Duke University, http://today.duke.edu/2005/09/quasicrystals.html

[4] Image: Inductiveload/Wikimedia Commons http://en.wikipedia.org/wiki/Penrose_tiling

[5] I.R. Fisher, Z. Islam, A.F. Panchula, K.O. Cheon, M.J. Kramer, P.C. Canfield, A.I. Goldman: Philos. Mag. **B 77** 1601–1615 (1998) Images: http://en.wikipedia.org/wiki/File:Zn-Mg-HoDiffraction.JPG

[6] P. Schweiss, W. Reichardt, M. Braden, G. Collin, G. Heger, H. Claus, A. Erb: Phys. Rev. **B 49**, 1387 – 1396 (1994)

B 2 The Structure of Complex Fluids [1]

G. Gompper
Institute of Complex Systems
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the $43^{\rm rd}$ IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Soft Matter Science investigates the structure and dynamics of macromolecules, macromolecular aggregates, and small particles in solution. The notion "Soft Matter" has been coined by Pierre-Gilles De Gennes, nobel prize winner 1991 for "his discovery that methods for the investigation of ordering phenomena in simple systems can be generalized to complex forms of matter, in particular for liquid crystals and polymers". The notion "Complex Fluids" is often used synonymously with Soft Matter.
The classical fields of Soft Matter Science are polymer solutions and melts, colloidal suspensions, and amphiphilic systems. The early days of the investigation of such systems dates back at least to the beginning of the last century. However, it has been recognized only in the 1980s and 1990s that these three classes of materials should really be seen as special cases of Soft Matter systems. The reasons for this realization is four-fold:

- All Soft Matter systems are characterized by a mesoscopic structural length scale, in the range of tens of nanometers to tens of micrometers.
- The systems are characterized by a typical energy scale comparable to the thermal energy k_BT (where k_B is Boltzmann's constant, and T is room temperature), so that thermal fluctuations play a large role.
- The constitutive macromolecules can have many architectures, which are intermediate between linear synthetic polymers, hard-sphere colloids, and short-chain amphiphilic molecules, as illustrated in Fig. 1.
- Many systems studied today contain components of different character, such as polymers and colloids, which is essential to construct new materials with tailored properties.

This chapter focuses on amphiphilic systems, because this is the archetypical example of structure formation in complex fluids. Here, the building blocks are amphiphilic molecules, which consist of a polar head and a non-polar hydrocarbon tail. Amphiphilic molecules are also often called "surfactants", as an abbreviation for **surf**ace **act**ive substance. Structure formation in mixtures with water and/or oil is then driven by the "hydrophobic effect" that polar and non-polar substances do not want to mix — a example from daily live is salad sauce, where the oil inevitably separates from the water after shaking or mixing. In oil-water-amphiphile mixtures, this leads to the formation of amphiphilic monolayers at the oil-water interface, and thereby to a reduction of the surface tension. In mixtures of water and amphiphilic molecules, bilayers form, in which the polar heads are directed towards the water and shield the hydrocarbon tails, which are buried inside the bilayer, from water contact. The resulting structures are shown schematically in Fig. 2. The formation of amphiphilic mono- or bilayers is only the first level of structure formation. On a second level, this aggregates are the building blocks of larger structures and mesophases. The understanding of the physical mechanisms of this structure formation, and how it can be revealed by scattering experiments, is the content of this chapter.
The behavior of other soft matter systems, in particular of polymer solutions and melts, will be explained in Chaps. B.6, E.2 and E.3.
The statistical physics of amphiphilic systems, membranes, and related subjects of structure formation in complex fluids, has been discussed in recent years in a many books and reviews, *inter alia* by Porte (1992) [3], Gompper & Schick (1994) [4], Gelbart et al. (1994) [5], Safran (1994) [6], Evans & Wennerström (1994) [7], Lipowsky and Sackmann [8], David et al. (1996)

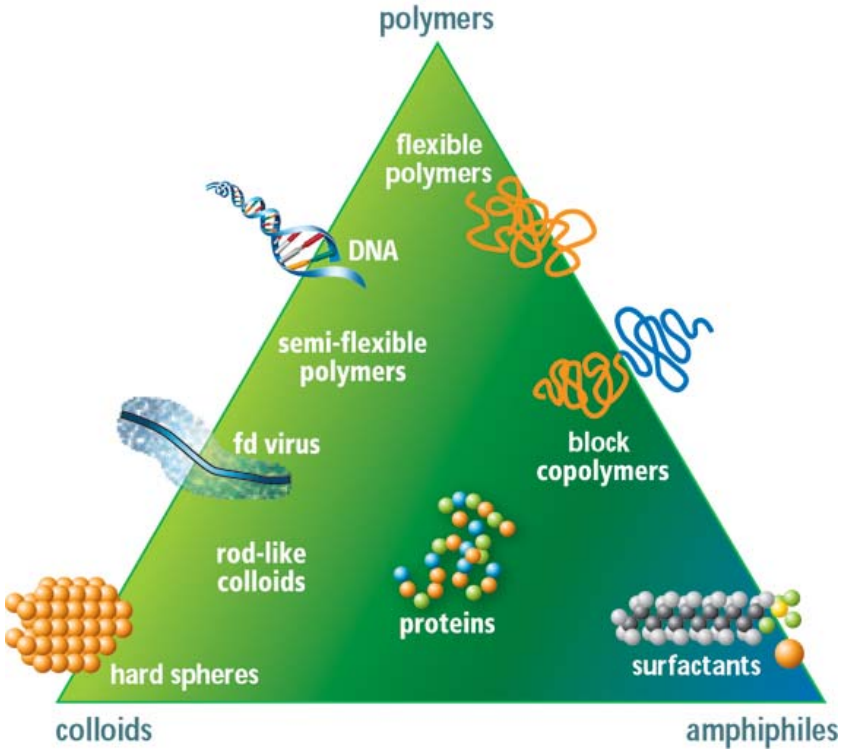


Fig. 1: *The "magic triangle" of soft matter, which shows that the classical fields of colloids, polymers and amphiphilic systems have merged into one. The aspect ratio (length/width) increases in the vertical direction, the amphiphilicity in the horizontal direction. From Refs. [1, 2].*

[9], Seifert (1997) [10], Gompper & Kroll (1997) [11], Cates & Evans (2000) [12], Dhont et al. (2002) [13], Nelson et al. (2004) [14], Witten & Pincus (2006) [15], Poon & Andelman (2006) [16], Dhont et al. (2008) [17], and Dhont et al. (2011) [18].

2 Oil-Water-Surfactant Mixtures

Phase Behavior of Water-Surfactant Mixtures — The experimental phase diagram of the binary mixture of the non-ionic surfactant $C_{12}E_5$ and water is shown in Fig. 3. The phase diagram is dominated by the two-phase coexistence between a dilute micellar phase L_1' and a more concentrated micellar phase L_1'' at low temperatures. Here, micelles are small spherical aggregates of amphiphilic molecules, with a hydrocarbon core and layer of polar heads outside, as shown schematically in Fig. 5 below. At high surfactant concentration, a lamellar phase L_α — a one-dimensional stack of surfactant bilayers separated by thin water films — is stable in this temperature range. At higher temperatures, the lamellar phase suddenly becomes stable already at very high dilution of only a few percent surfactant. Simultaneously, a sponge phase L_3 appears at even lower surfactant concentration.

Both the lamellar phase and the sponge phase are made of surfactant *bilayers*. Since the two leaflets of the bilayers are equivalent, these membranes cannot have a preferred curvature by symmetry (in contrast to monolayers, where in general a preferred curvature towards the oil- or

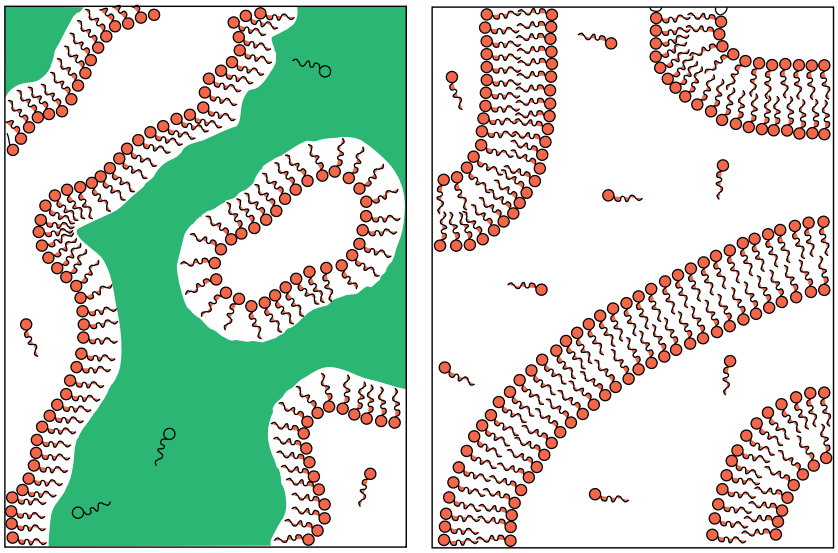


Fig. 2: *Amphiphilic molecules self-assemble into complex aggregates due to the hydrophobic effect. In mixtures with oil and water, they form monolayers at the water-oil interface. In mixtures with water, they form bilayers.*

water-side exists).

Lamellar Phase L_α **—** The stability of the lamellar phase at these strikingly low surfactant concentrations has caught the imagination of physicists in the late seventies of the last century. It was proposed by Helfrich (1978) [20] that this swelling behavior of the lamellar phase under the addition of water can be understood by the repulsive force generated by the suppression of long wave-length undulation modes of an individual membrane by its neighbors in the stack. This leads to a reduction of its entropy, and therefore to an increase of the free energy Δf per unit area, which was predicted to have the form [20]

$$\Delta f = c_\infty \frac{(k_B T)^2}{\kappa} d^{-2} \tag{1}$$

where d is the repeat distance of the lamellar phase, and κ is the bending rigidity of the bilayer. This behavior has since been investigated and verified experimentally (Safinya et al. 1986, 1989) [21, 22]. Monte Carlo simulations [23] give a value $c_\infty = 0.106$ for the value of the universal coefficient in Eq. (1).

Sponge Phase L_3 **—** Freeze-fracture microscopy reveals the mesoscopic structure of the sponge phase, as shown in Fig. 4. The fluid is frozen very rapidly to preserve its mesoscopic structure during the freezing process, then fractured into pieces. Fortunately, the fracture surface follows preferentially the midplane of the bilayer. Therefore, the three-dimensional structure of the membrane is revealed by this technique. Fig. 4 demonstrates impressively that the sponge phase is a complicated network of water channels, which are separated by membranes with a locally *saddle-shaped* structure.

The "sponge phase" has gotten it's name because its structure resembles the structure of a *sea-water* sponge (compare Fig. 4). The pore space of the sea-water sponge has to be connected in order for the water to be able to penetrate the sponge; its solid component has also to be connected to keep the sponge together. Thus, both objects are "bicontinuous" (see below).

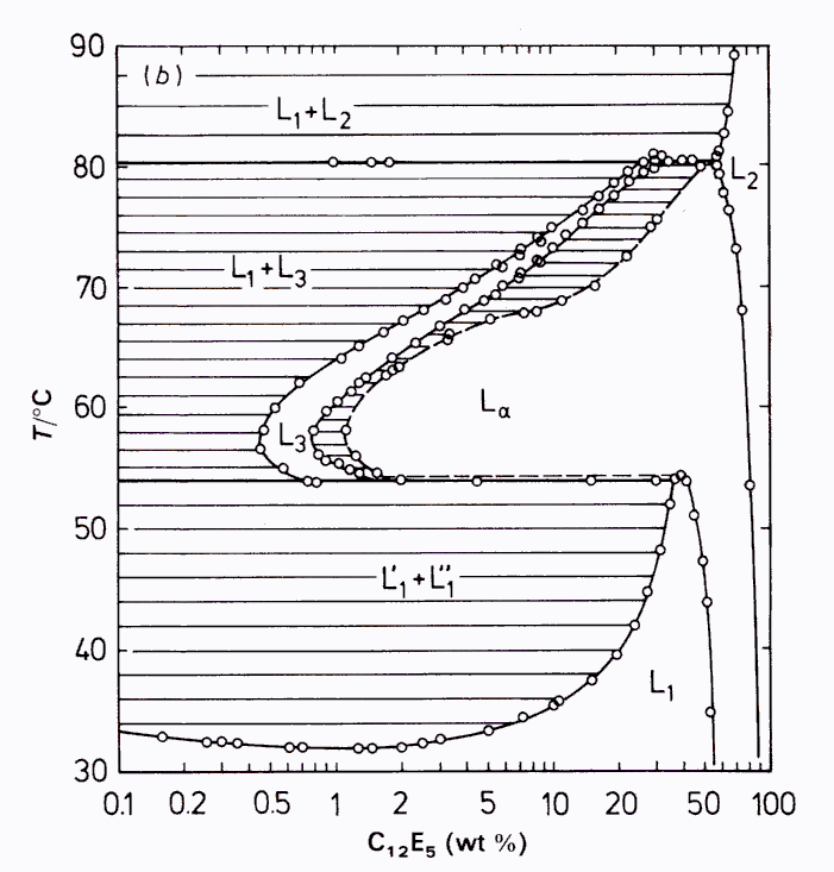


Fig. 3: *Experimental phase diagram of the binary mixture of $C_{12}E_5$ and water (Strey et al. 1990) [19]. The phase diagram is dominated by the two-phase coexistence between a dilute (L_1') and a more concentrated (L_1'') micellar phase at low temperatures. At higher temperatures, the lamellar phase (L_α) becomes stable already at high dilution of only a few percent surfactant. Simultaneously, a sponge (L_3) phase appears at even lower surfactant concentration. The phase at high surfactant concentration (L_2) is an inverted micellar phase (with surfactant tails pointing outwards). The hatched areas are two-phase coexistence regions. Note the logarithmic scale of the abscissa.*

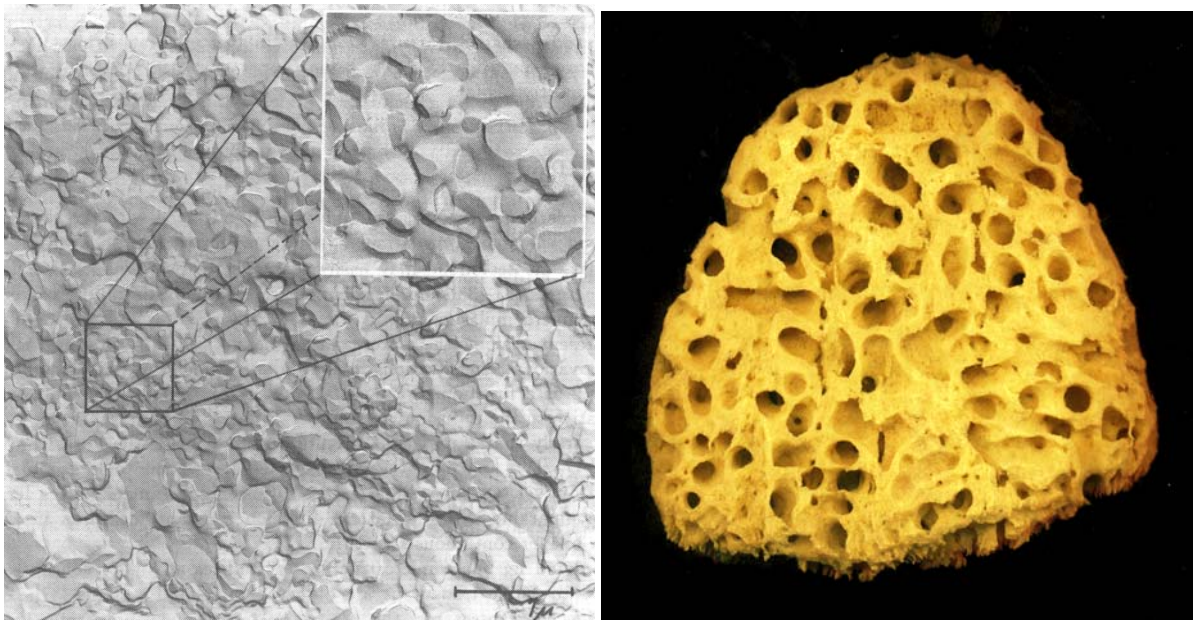


Fig. 4: *Left: Freeze-fracture-microscopy picture of a sponge phase in the system water/NaCl and AOT (Strey et al. 1992) [24]. Note the local saddle-shape structure of the membrane, which can be seen most clearly in the inset. Right: A sea-water sponge.*

However, the length scales are very different.

Microemulsions — When oil is added as a third component to a mixture of water and surfactant, the phase diagram obviously gets considerable more complicated [4], as shown schematically in Fig. 5. The spontaneous curvature of the monolayer membrane is in general non-zero in these systems. For most non-ionic surfactants, the spontaneous curvature depends sensitively and approximately linearly on temperature (compare Sec. 5), with curvature towards the oil-side at lower temperatures and curvature towards the water-side at higher temperatures (this can be understood by the diminishing hydration shell of the polar heads with increasing temperature). Therefore, the spontaneous curvature vanishes for one particular temperature, the hydrophobic-hydrophilic-balance temperature $\bar{T}$. In a temperature interval around $\bar{T}$, a phase is stable, which is macroscopically homogeneous and isotropic, and which contains equal amounts of oil and water. This phase is called a *microemulsion* . It has a very similar mesoscopic structure as the sponge phase:

- The membrane in the sponge is a bilayer, in the microemulsion it is a monolayer.
- The microemulsion consists of network of oil- and water-channels, which are separated by a surfactant monolayer. The sponge phase consists of (at least) two distinct networks of water-channels, which are separated by a bilayer.

A microemulsion is called *bicontinuous*, because both the oil- and the water-networks span the whole system. Since the sponge phase has an equivalent mesoscopic structure with two networks of water channels, it is also called bicontinuous.

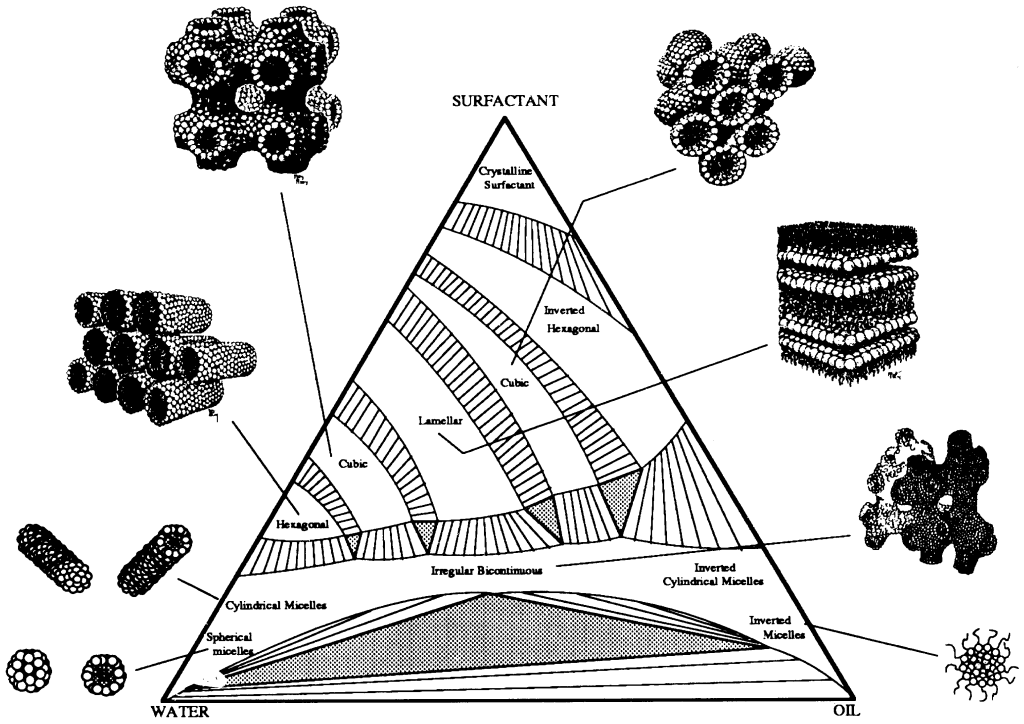


Fig. 5: *Schematic phase diagram of ternary amphiphilic systems at fixed temperature. From Davis et al. (1987) [25].*

3 Theoretical Approaches

3.1 Length Scales

A theoretical model always depends on the type of phenomena to be studied, and on the basic length scales to be described. This is not much different from experimental studies, where different techniques are appropriate for different ranges of length scales. For oil-water-surfactant mixtures, the different levels of descriptions, the appropriate length scales, and the the corresponding degrees of freedom, are illustrated in Fig. 6. On the molecular level, we are dealing with a three-component mixture of different molecules. The length scale is therefore atomistic, and the degrees of freedom are the atomic coordinates. This is a lot of information, which is only partially useful for a structured fluid with large oil-rich and water-rich domains. Therefore, we can zoom out a bit and only consider the concentrations fields, which are averaged over the local molecular conformations and orientations. On this level, we can already see that surfactant assembles at the mesoscopic oil-water interface. However, the domain structure is still difficult to describe. Thus, if we are interested in the shape and topology of the interfaces, we have to step back even further, and describe the membrane as a mathematical surface, with its shape and fluctuations controlled by curvature elasticity.

For the calculation of phase behavior and scattering intensities, the descriptions of the concentration level and on the membrane level are particularly useful.

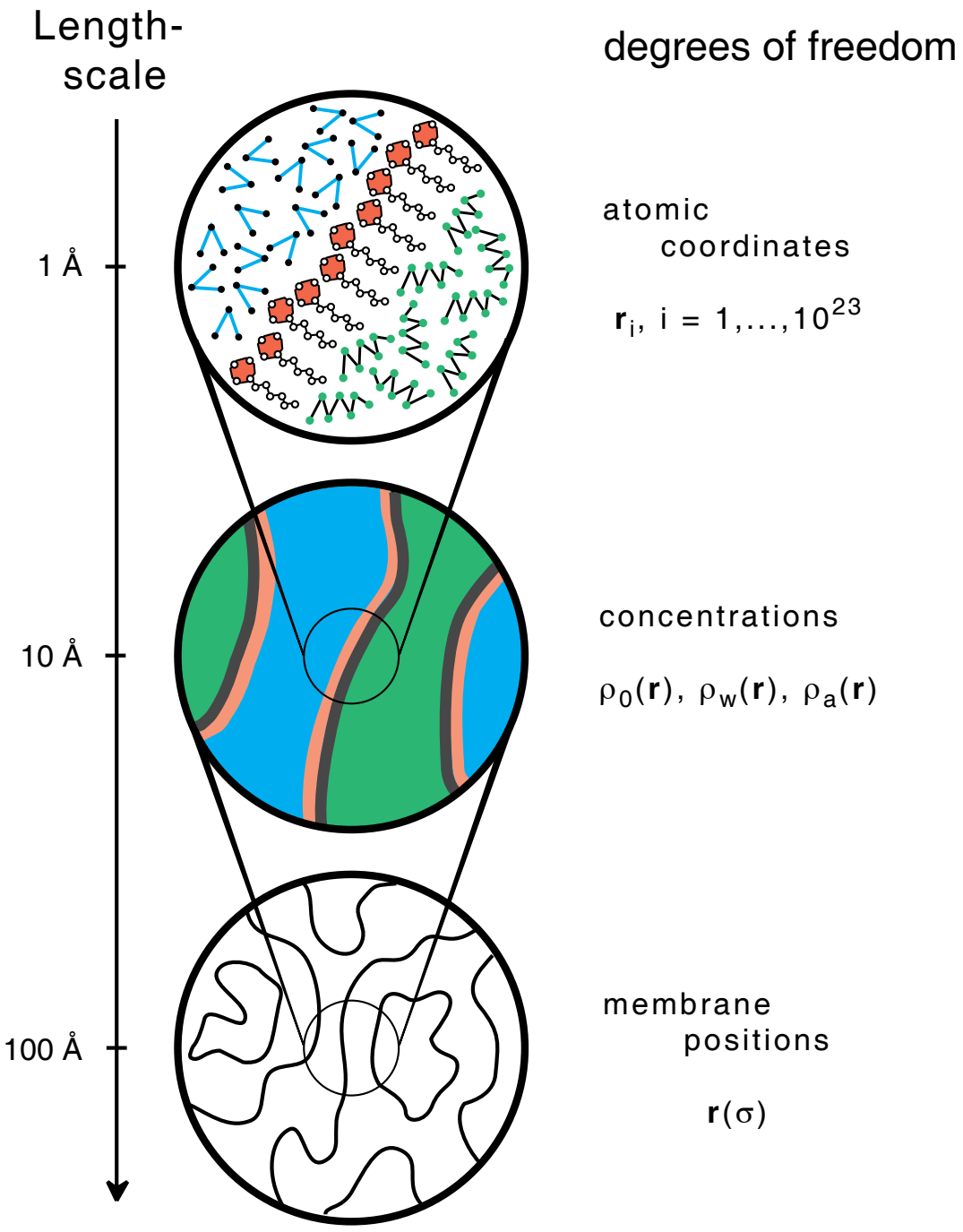


Fig. 6: *Modeling microemulsion structure on different length scales. The basic degrees of freedom on each level of coarse graining are indicated.*

3.2 Membrane Model

Curvature Energy and Fluctuating Surfaces — We want to try to understand the structure and phase behavior of binary mixtures of water and surfactant on the basis of a model, in which the surfactant film is described by a surface, whose shapes and fluctuations are controlled by the curvature energy (2). First attempts for such a description go back to Helfrich (1973) [26], Scriven (1976) [27], de Gennes and Taupin (1982) [28], and Safran et al. (1986) [29].
The energy functional of a *fluid membrane* must satisfy the following conditions

- Motion invariance:
  The energy must be invariant under translations and rotations of the whole membrane
- Reparametrization invariance:
  Since the molecules can diffuse freely within the membrane, no preferred coordinate system can exist — which would correspond physically to a labeling of distinct constitutive elements of the membrane. Therefore, the energy must be invariant under a change of the coordinate system.

These conditions imply that in an expansion in powers of the inverse radii of curvature, the curvature energy to leading order is given by [30, 26]

$$\mathcal{H}_{curv} = \int dS \left[\gamma + 2\kappa(H - c_0)^2 + \bar{\kappa}K + ...\right] \tag{2}$$

where the integration is over the whole membrane surface,

$$2H = c_1 + c_2 \quad , \quad K = c_1 c_2 \tag{3}$$

are the mean and Gaussian curvatures, respectively, which are expressed in terms of the principal curvatures $c_1 = 1/R_1$ and $c_2 = 1/R_2$ at each point of the membrane, compare Fig. 7. It is interesting to note that H and K are the trace and determinant of the curvature tensor, and therefore satisfy the invariance principles stated above. The parameters γ, κ, $\bar{\kappa}$ and c_0 are the elastic constants of the membrane, which depend on the structure and interactions of the constitutive molecules. Their physical meaning is:

- γ is the surface tension,
- κ is the bending rigidity,
- c_0 is the spontaneous (or preferred) curvature,
- $\bar{\kappa}$ is the saddle-splay modulus – relevant only for topology changes.

From a statistical mechanics point of view, the calculation of the thermodynamic properties of an ensemble of fluctuating surfaces is a formidable problem. It amounts to calculating the partition function

$$Z = \sum_{topologies} \int' \mathcal{D}\mathbf{R}(\underline{\sigma}) \exp\{\mathcal{H}[\mathbf{R}(\underline{\sigma})]/k_B T\} \tag{4}$$

where $\int' \mathcal{D}\mathbf{R}(\underline{\sigma})$ denotes an integration over all possible shapes with parametrization $\mathbf{R}(\underline{\sigma})$ of the surface at fixed topology, where $\underline{\sigma}$ is a two-dimensional coordinate system on the surface. However, this integral cannot be just over all possible parametrizations $\mathbf{R}(\underline{\sigma})$ of a surface of

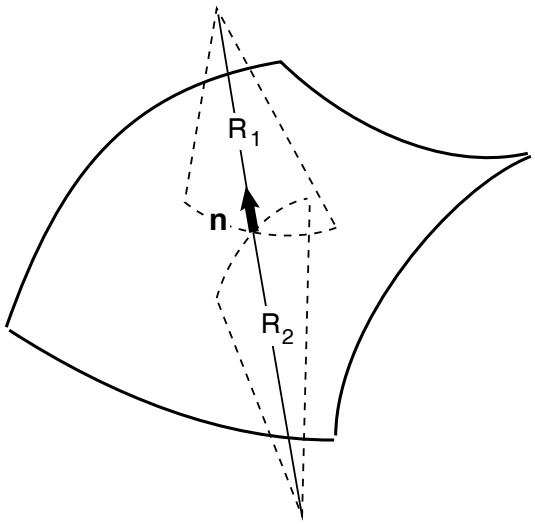


Fig. 7: *The two principal radii of curvature, R_1 and R_2, of a point on the surface.*

fixed topology, but has to be restricted to those parametrizations, which lead to physically different shapes in the embedding space; this is indicated by the prime. Finally, the contributions off all different topologies have to be summed over. It is clear that this problem is sufficiently complex that no exact solution will be found anytime soon. Therefore, approximations have to be made in order to get some insight into the behavior of these phases.

Energy Considerations — Let us first take a closer look at the curvature *energy* itself, and consider the effect of fluctuations in a second step. It is easily seen that for $c_0 = 0$ the curvature Hamiltonian (2) can be rewritten in the form

$$E = \int dS \left[\frac{1}{2}\kappa_+ \left(\frac{1}{R_1} + \frac{1}{R_2} \right)^2 + \frac{1}{2}\kappa_- \left(\frac{1}{R_1} - \frac{1}{R_2} \right)^2 \right] \tag{5}$$

with

$$\kappa_+ = \kappa + \frac{1}{2}\bar{\kappa} \quad , \quad \kappa_- = -\frac{1}{2}\bar{\kappa} \tag{6}$$

This form is convenient to show that the stability of lamellar phase requires

$$\kappa_+ > 0 \ , \ \kappa_- > 0 \tag{7}$$

since otherwise the energy of the system could be lowered without bounds by making the term with negative amplitude very large in magnitude. With Eq. (6), the stability conditions (7) are equivalent to

$$-2\kappa < \bar{\kappa} < 0 \tag{8}$$

Instabilities occur as κ_+ or κ_- approach zero:

- For $\kappa_+ \to 0$, it costs very little energy to make $|c_1 + c_2|$ large, as long as $|c_1 - c_2|$ remains small. Thus, in this limit it costs a very small energy to form small spherical objects with $c_1 \simeq c_2$, which in the case of monolayers are called micelles, in the case of of bilayers *vesicles*.

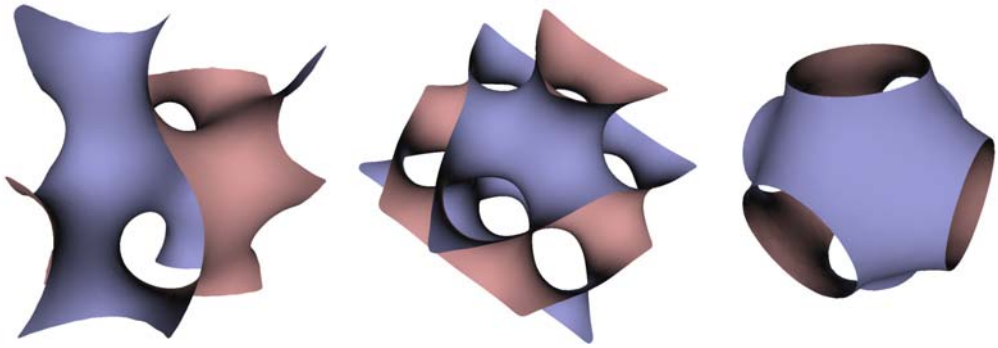


Fig. 8: *The three most important triply periodic minimal surfaces (TPMS) of cubic symmetry: Schwarz P-surface (right), Schwarz D-surface (middle), and Schoen's gyroid surface (left) [31].*

- For $\kappa_- \to 0$, it costs very little energy to make $|c_1 - c_2|$ large, as long as $|c_1 + c_2|$ remains small. Thus, in this limit it costs a very small energy to form surfaces with $c_1 \simeq -c_2$ and small domain size, such as microemulsions or sponge phases.

It is interesting to note that sufaces with $c_1 = -c_2$ have been studied extensively for more than 100 years. These surfaces are called *minimal surfaces*. In the context of amphiphilic systems, triply periodic minimal surfaces (TPMS) of cubic symmetry, shown in Fig. 8, are particularly relevant. These phases are also often called *plumber's nightmare phases*, because this multiply connected system of tubes would be almost impossible for a plumber to prevent such a structure from leaking.

Scaling Considerations — Let us consider an arbitrarily shaped piece of membrane with fixed geometry. Let the local curvatures be denoted by c_i, and the curvature energy of this patch be E_b. When this membrane piece is rescaled uniformly by a scale factor λ — i.e. the "shape" remains the same but the size changes — then the new local curvatures become $c_i' = \lambda^{-1} c_i$, and the curvature energy changes to

$$\begin{aligned} E_b' &= \int dS' \left\{ 2\kappa H'^2 + \bar{\kappa} K' \right\} \\ &= \int dS \lambda^2 \left\{ 2\kappa \lambda^{-2} H^2 + \bar{\kappa} \lambda^{-2} K \right\} \equiv E_b \end{aligned} \tag{9}$$

Thus, the curvature energy remains invariant under uniform scale transformations. This results implies immediately that the energy *density*, which is the energy *per unit volume*, depends on the surface-to-volume ratio, S/V, which is proportional to the amphiphile concentration, as [32]

$$E_b/V \sim (S/V)^3 \tag{10}$$

Eq. (10) shows that the curvature energies of all membrane structures scale in exactly the same way (although with different prefactors), so that an intersection of the energy curves for different structures as a function of S/V cannot occur. This result allows an important conclusion. On the basis of the curvature energy alone, without thermal fluctuations, we will not be able to understand transitions between different mesophases as the amphiphile concentration is varied, such as the lamellar-to-sponge transition in Fig. 3 at fixed temperature. It must be the *entropy*

which is responsible for these transitions.

Membrane Fluctuations — The thermal fluctuations on small scales — i.e. small compared to the typical domain size of a sponge phase, or small compared to the average separation of membranes in a lamellar phase — modify the curvature elasticity on larger scales. This effect can be easily demonstrated with a sheet of paper for *polymerized* membranes (with fixed connectivity of neighboring molecules): A corrugated sheet has a higher bending rigidity on scales larger than the characteristic ripples than a smooth sheet. Thus, "fluctuations" increase the rigidity in this case. In *fluid* membranes, fluctuations have the opposite effect of softening the rigidity on larger scales. There is no macroscopic example to demonstrate this effect. It follows from renormalization group calculations [33, 34], which show that thermal fluctuations lead to renormalized, *scale-dependent* rigidities κ^R, $\bar{\kappa}^R$ of the form

$$\kappa_R(\xi) = \kappa - \alpha \frac{k_B T}{4\pi} \ln(\xi/a) \tag{11}$$

$$\bar{\kappa}_R(\xi) = \bar{\kappa} - \bar{\alpha} \frac{k_B T}{4\pi} \ln(\xi/a) \tag{12}$$

with $\alpha = 3$ and $\bar{\alpha} = -10/3$ (Peliti & Leibler 1985, David 1989) [33, 35]. This implies immediately that the elastic moduli of Eq. (5) are also renormalized, and are given by

$$\kappa_\pm^R(\xi) = \kappa_\pm - \alpha_\pm \frac{k_B T}{4\pi} \ln(\xi/a) \tag{13}$$

with

$$\alpha_+ = \alpha + \frac{1}{2}\bar{\alpha} = \frac{4}{3} \quad , \quad \alpha_- = -\frac{1}{2}\bar{\alpha} = \frac{5}{3} \tag{14}$$

Phase Behavior — When the fluctuation on short scales are taken into account, the instabilities of the lamellar phase discussed above now occur where the *renormalized* rigidities vanish, i.e. at $\kappa_\pm^R(\xi) = 0$. The length scale ξ, where the renormalization is cut off, is determined by the average separation of neighboring membranes, which is given by the membrane volume fraction

$$\Psi \equiv \frac{aS}{V} = a/\xi \tag{15}$$

where a is again a molecular scale, here the length of the amphiphiles. This finally implies (Morse 1994) [36]

$$\ln(\Psi) = -\frac{4\pi}{\alpha_\pm} \frac{\kappa_\pm}{k_B T} \tag{16}$$

The phase diagram, which is predicted on the basis of this calculation, is shown in Fig. 9. This phase diagram does not only contain the instability lines (16), but also results of a somewhat more detailed calculation, in which estimates of the free energies of the sponge phase (based on an approximation by minimal surfaces, so that the free energy is taken to be $\bar{\kappa}_R(\xi) \int dS\ K$) and of the lamellar phase (based on the steric repulsion expression (1)) are compared. Fig. 9 shows that the two results are in good agreement.

The phase diagram is dominated by a V-shaped region, where the lamellar phase is stable. With decreasing amphiphile concentration, a phase transition to the sponge phase occurs for $0 > \bar{\kappa}/\kappa > -10/9$, while a transition to a vesicle phase occurs for $-10/9 > \bar{\kappa}/\kappa > -2$. The sponge phase exists only in a narrow strip parallel to the instability line. When the amphiphile concentration is decreased further, *emulsification failure* occurs, i.e. the sponge phase cannot take up any more water, and therefore coexists with an (almost pure) water phase at these low membrane volume fractions.

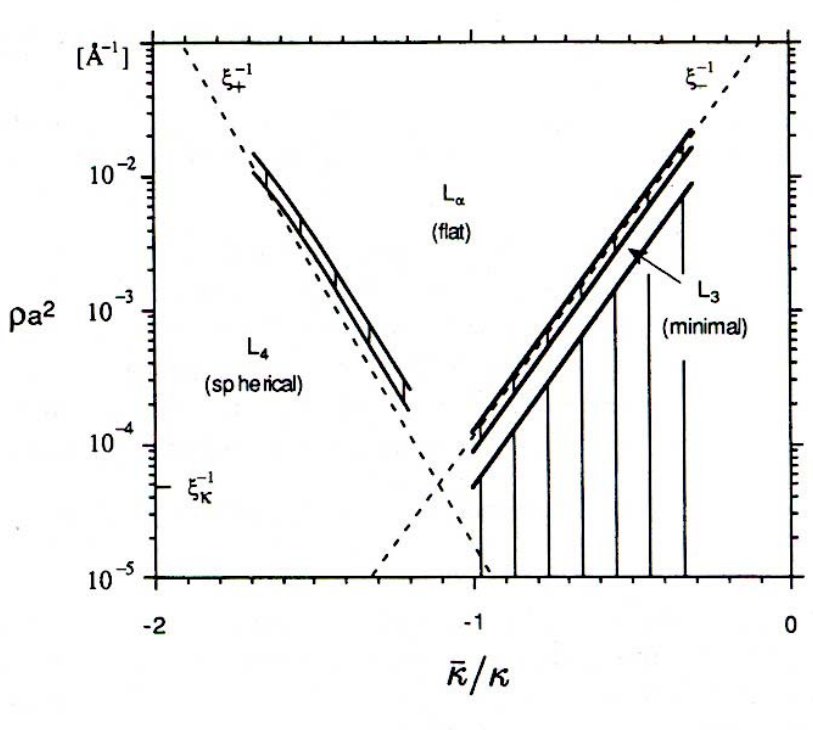


Fig. 9: *Phase diagram of binary or balanced ternary amphiphile mixtures, as predicted on the basis of the renormalization of the bending rigidity κ and the saddle-splay modulus $\bar{\kappa}$, for fixed $\kappa/k_BT = 2$ (Morse 1994) [36]. The phase behavior is shown as a function of the membrane volume fraction Ψ (denoted here ρa^2) and the saddle-splay modulus $\bar{\kappa}$. L_3 denotes the sponge phase, L_4 a vesicle phase. The two dashed lines are the stability boundaries (16), with the left line corresponding to $\kappa_+^R = 0$, the right line to $\kappa_-^R = 0$.*

3.3 Ginzburg-Landau Model

An alternative description of microemulsion starts from the point of view of ternary mixtures with three spatially varying concentration fields $\rho_o(\mathbf{r})$, $\rho_w(\mathbf{r})$, and $\rho_s(\mathbf{r})$ for oil, water, and surfactant, respectively. Such a description has been used extensively for near-critical binary mixtures. For incompressible of, say, oil and water, $\rho_o(\mathbf{r}) + \rho_w(\mathbf{r})$ =const., so that only the local concentration difference $\phi(\mathbf{r}) \equiv \rho_o(\mathbf{r}) - \rho_w(\mathbf{r})$, which is denoted "order parameter field", has to be considered. The model is then defined by the free energy functional

$$\mathcal{F}[\phi] = \frac{1}{2}\int d^3r \left[b_0(\nabla\phi(\mathbf{r}))^2 + f(\phi)\right] \tag{17}$$

Here, the free energy density of homogeneous order parameter fields has two minima below the critical temperature T_c, corresponding to the order parameters of the coexisting oil-rich and water-rich phases, and is usually written in an expansion for small order parameters as $f(\phi) = \tau\phi^2 + g\phi^4$, where $\tau = (T - T_c)/T_c < 0$ measures the distance from the critical point. Above T_c, where $\tau > 0$, $f(\phi)$ has a single minimum at $\phi = 0$ (completely mixed state). The first term on the right-hand side of Eq. (17), with $b_0 > 0$, penalizes inhomogeneous order-parameter configurations, and is therefore closely related to the oil-water interface tension.
This approach can be generalized to ternary amphiphilic systems [37]. In the simplest case, where the surfactant molecules are assumed to mostly populate the microscopic oil-water interface, the surfactant density does not have to be taken into account as an additional order parameter. Instead, it modifies only the structure of the free-energy functional, which now becomes [37]

$$\mathcal{F}[\phi] = \frac{1}{2}\int d^3r \left[c_0(\nabla^2\phi(\mathbf{r}))^2 + b(\phi)(\nabla\phi(\mathbf{r}))^2 + f(\phi)\right] \tag{18}$$

Here, the function $f(\phi)$ has now three minima, corresponding to the coexisting oil-rich, water-rich and microemulsion phases. The function $b(\phi)$ has a single minimum in the microemulsion phase, with $b_0 = b(0) < 0$; this negative value of the gradient-term in Eq. (18) corresponds to the fact that the system gains free energy by increasing the amount of oil-water interface area. This growth is limited by the first term in Eq. (18), with $c_0 > 0$, which is closely related to the bending rigidity of the surfactant layer.
In the microemulsion phase, where $\langle\phi(\mathbf{r})\rangle = 0$, the free-energy functional (18) can be expanded to quadratic order in the order-parameter fluctuations,

$$\mathcal{F}[\phi] = \frac{1}{2}\int d^3r \left[c_0(\nabla^2\phi(\mathbf{r}))^2 + b_0(\nabla\phi(\mathbf{r}))^2 + a_0\phi^2\right] \tag{19}$$

In this approximation, the scattering intensity can be calculated exactly, as explained in Appendix A, with the result [38]

$$I(q) = \frac{1}{c_0q^4 + b_0q^2 + a_0} \tag{20}$$

The negative value of b_0 implies that there is a scattering peak at non-zero wave vector. It was shown already more than 20 years ago by Teubner & Strey (1987) [38] that the functional form of Eq. (20) indeed fits experimental scattering data very well, compare Fig. 10.
The correlation function in real space is obtained from the scattering intensity by Fourier transformation, as explained in detail in Chap. A.5. Fortunately, the form of Eq. (20) is simple enough that this Fourier transformation can be done analytically, so that [38, 4]

$$G(r) = \int \frac{d^2q}{(2\pi)^2} I(q) e^{-i\mathbf{q}\cdot\mathbf{r}} \sim \frac{1}{r} e^{-r/\xi} \sin k_0 r \tag{21}$$

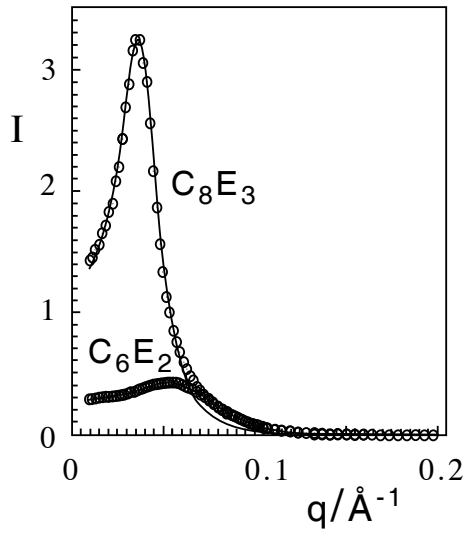


Fig. 10: *The scattering intensity $I(q)$ of a balanced microemulsion in bulk contrast for two different surfactants. The data are very well fitted by the functional form of Eq. (20) (Teubner & Strey 1987) [38].*

with *characteristic wave vector* k_0 and *correlation length* ξ, where

$$k_0 = \frac{1}{2}\left[2\sqrt{a_0/c_0} - b_0/c_0\right]^{1/2} \quad , \quad \xi^{-1} = \frac{1}{2}\left[2\sqrt{a_0/c_0} + b_0/c_0\right]^{1/2} \tag{22}$$

3.4 Gaussian Random Fields

A description of microemulsions and sponge phases, which is closely related to the Ginzburg-Landau approach, is to employ *level surfaces* of a scalar field $\Phi(\mathbf{r})$, with $\mathbf{r} \in \mathbb{R}^3$, which are defined by

$$\Phi(\mathbf{r}) = \alpha \tag{23}$$

This implicitly defines a surface in three-dimensional space. An illustration of this concept is shown in Fig. 11, where the level "surfaces" are shown for a scalar field in two dimensions. Thermal fluctuations of the scalar field $\Phi(\mathbf{r})$ — according to a Boltzmann weight $\exp(-\mathcal{H}_0)$ — then imply fluctuations of the level surfaces. The properties of the membranes are therefore induced by the statistical mechanics of the scalar field Φ, as they are in the Ginzburg-Landau model.
A particularly useful case are *Gaussian random fields* (GRF) , where the fluctuations of Φ are controlled by a Boltzmann weight with a *quadratic* Hamiltonian

$$\mathcal{H}_0 = \frac{1}{2}\int d^3r \Phi(\mathbf{r})\nu^{-1}(|\mathbf{r}-\mathbf{r}'|)\Phi(\mathbf{r}') \tag{24}$$

with *spectral density* $\nu(\mathbf{r}-\mathbf{r}')$. Requirements on $\nu^{-1}(|\mathbf{r}-\mathbf{r}'|)$ are that it decays sufficiently fast for large $|\mathbf{r}-\mathbf{r}'|$ so that the integral and the second and fourth moments are finite. The normalization

$$\langle\Phi(\mathbf{r})\rangle = 0 \quad , \quad \langle\Phi(\mathbf{r})^2\rangle = 1 \tag{25}$$

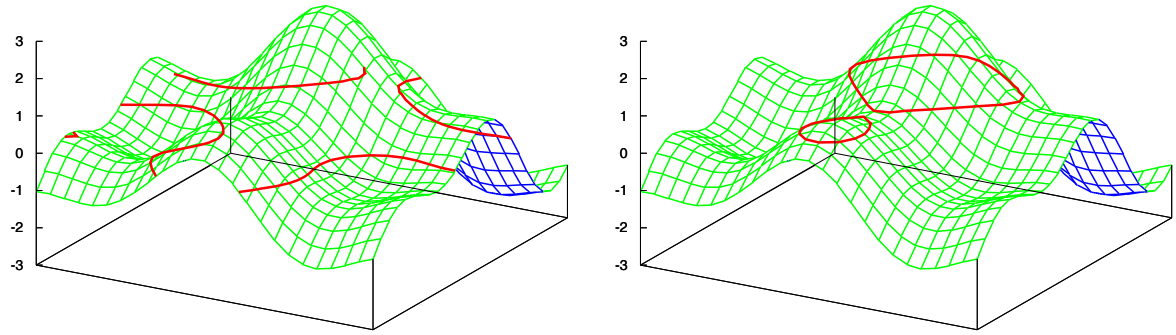


Fig. 11: *Level "surfaces" of a scalar field $\Phi(\mathbf{r})$, with $\langle\Phi(\mathbf{r})\rangle = 0$. For illustration purposes, $\mathbf{r} \in \mathbb{R}^2$, so that the level "surfaces" are in fact lines in this case. Left: Level surface (red lines) for $\alpha = 0$, so that the surface has no preferred curvature, $\langle H\rangle = 0$, and a sponge-like structure can be envisaged. Right: Level surface (red lines) for $\alpha = 1$, so that the level surface bends preferentially to one side, and droplet-like structures can be seen.*

is usually employed, where the first identity follows immediately from the $\Phi \to -\Phi$ symmetry of the Hamiltonian, while the second identity implies the normalization $\int d^3k\ \nu(k) = (2\pi)^3$. The advantage of the Gaussian random field model is that all averages can be calculated exactly. In particular, the *average geometry* of the level surfaces can be calculated for Gaussian random fields (GRF), with the analytical results (Teubner 1991) [39]:

$$S/V = \frac{2}{\pi}\exp[-\alpha^2/2]\sqrt{\frac{1}{3}\langle q^2\rangle_\nu} \tag{26}$$

$$\langle K\rangle = -\frac{1}{6}\langle q^2\rangle_\nu(1-\alpha^2)\,, \tag{27}$$

$$\langle H\rangle = \frac{1}{2}\alpha\sqrt{\frac{\pi}{6}\langle q^2\rangle_\nu}\,, \tag{28}$$

$$\langle H^2\rangle = \langle K\rangle + \frac{1}{5}\frac{\langle q^4\rangle_\nu}{\langle q^2\rangle_\nu} \tag{29}$$

where

$$\langle q^n\rangle_\nu \equiv \int \frac{d^3q}{(2\pi)^3} q^n \nu(q) \tag{30}$$

The calculation of the surface density S/V is sketched in Appendix B to illustrate the calculation of geometrical averages for Gaussian random fields.
The results (26) to (29) imply some interesting conclusions for the symmetric case $\alpha = 0$:

- Eq. (28) shows that mean curvature vanishes, $\langle H\rangle \equiv 0$, so that $\alpha = 0$ corresponds to a balanced system, such as in sponge phases or balanced microemulsions. Compare Fig. 11.

- From Eq. (27), we find $\langle K\rangle = -\frac{1}{6}\langle q^2\rangle_\nu < 0$; this implies that the average geometry of level surfaces in Gaussian random fields is *saddle-shaped*. Eqs. (27) and (26) can be combined to give the dimensionless quantity $\langle K\rangle(S/V)^{-2} = -\pi^2/8 = -1.234$, which

characterizes the topology of a characteristic "unit cell" of the sponge phase. It is important to note that this number is *universal* in the Gaussian random field model, since it is completely independent of the spectral density ν.

3.5 Variational Approach

In Sec. 3.4, the spectral density ν has never been specified. The application of the Gaussian random field model to sponge phases therefore remains qualitative on this level [40, 41, 39]. What is missing is a relation of the spectral density ν to the Hamiltonian which describes the system under consideration. For a membrane ensemble controlled by the curvature energy, considerable progress can be made by employing the Gaussian random field model in a variational approach (Pieruschka & Safran 1993) [42].
The variational approach is based on the Gibbs-Bogoliubov-Feynman inequality [43, 44]. It relates the energies and free energies of the system under consideration (in our case the membrane ensemble with curvature energy) to a reference system, which can be treated more easily (in our case the Gaussian random field model).
Let the full system be described by a Hamiltonian $\mathcal{H}$ and have free energy F, and the reference system be described by Hamiltonian $\mathcal{H}_0$ and have free energy F_0. The Gibbs-Bogoliubov-Feynman inequality is based on the inequality $\langle \exp(f) \rangle \geq \exp\langle f \rangle$ for any function $f \in \mathbb{R}$ [44]. This implies

$$F \leq F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 \tag{31}$$

The main idea of the variational approach is then to choose a reference system with free parameters, which can be used to find an optimal (i.e. lowest) upper bound for the free energy of the full system.
The application of this approach to membrane ensembles requires the calculation of $\langle \mathcal{H} \rangle_0$ and F_0. The average curvature energy is

$$\langle \mathcal{H}_{curv} \rangle_0 = S \left(2\kappa \langle H^2 \rangle_0 + \bar{\kappa} \langle K \rangle_0 \right) \tag{32}$$

which can be easily expressed as a functional of the spectral density with Eqs. (26)-(29). The calculation of the free energy of a Gaussian model is a standard problem of statistical field theory, as described for example in detail in Ref. [44], and has the result

$$F_0 = -\frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \ln(\nu(k)) \tag{33}$$

Functional differentiation of Eq. (31) with Eqs. (32) and (33) then gives the *optimal spectral density* (Pieruschka & Safran 1995) [45]

$$\nu(q) = \frac{1}{cq^4 + bq^2 + a} \tag{34}$$

where a, b, and c are now functions of κ and S/V.
The spectral density (34) can be identified with the scattering intensity $I(q)$ in bulk contrast of ternary microemulsions, i.e. with scattering contrast between oil and water, in the same way as discussed for the Ginzburg-Landau model in Sec. 3.3 above.
With the use of Eq. (22), the results of the variational approach then imply for the characteristic length scales in the correlation function (for large κ/k_BT):

$$k_0 \sim S/V \quad , \quad \xi \sim \kappa (S/V)^{-1} \quad , \quad k_0 \xi \sim \kappa \tag{35}$$

Thus, the wave vector k_0 increases linearly with the amphiphile concentration, while the correlation length ξ decreases. However, the dimensionless ratio $k_0\xi$ of the two length scales, which characterize a microemulsion or sponge phase, is predicted to be *independent* of the amphiphile concentration, and only to depend on the bending rigidity. These are very specific predictions, which can be tested experimentally and by simulations.

3.6 Comparison of Theoretical Approaches

We have now predictions by two different theoretical models, which we abbreviate by the names 'Fluctuating Membranes" and "Gaussian Random Fields". It is instructive to summarize the main points discussed above:

- "Fluctuating Membranes" and "Gaussian Random Fields" are *complementary models*
- The variational approach provides a link between these models
- A direct comparison is possible for the *free energy*:
  Fluctuating Membranes —

$$\begin{aligned} f = F/V &\sim [\kappa_R(\Psi) + \bar{\kappa}_R(\Psi)]\Psi^3 \\ &= [A + B\ln(\Psi)]\Psi^3 \end{aligned} \tag{36}$$

  as shown in Refs. [46, 47, 32, 3].
  Gaussian Random Fields —

$$f = F/V = A'\Psi^3 + B'\ln(\Psi) \tag{37}$$

  as derived in Ref. [45]. It is obvious from the comparison of these two equations that although the results look similar they do *not* agree!

4 Monte Carlo Simulations of Dynamically Triangulated Surfaces

As in all entropically dominated, strongly fluctuating systems, simulation methods provide a very important tool to obtain information about the system properties, which are not easily accessible by purely analytical approaches. For sufficiently small systems, thermal averages can be calculated by simulation techniques with very high precision. The extrapolation to large system sizes is sometimes difficult, but often possible with good reliability.

Motivation — What are the particular reasons to apply simulations to membrane ensembles:

- Clarify the discrepancy between functional forms of the free energies of the "Fluctuating Membranes" and the "Gaussian Random Field" approaches, Eqs. (36) and (37).
- Different results have been obtained for the universal values of the prefactors α and $\bar{\alpha}$ of the logarithmic renormalization of κ and $\bar{\kappa}$ — compare Eqs. (11) and (12), respectively:
  $\alpha = +3$, $\bar{\alpha} = 10/3$ (Peliti & Leibler 1985; Cai *et al.* 1994) [33, 34];
  $\alpha = +1$, $\bar{\alpha} = 0$ (Helfrich 1985) [48]
  $\alpha = -1$, $\bar{\alpha} = 0$ (Helfrich 1998) [49]
  The phase diagram of Fig. 9 has been calculated for $\alpha = +3$, $\bar{\alpha} = 10/3$. It looks very different for the other values given above.

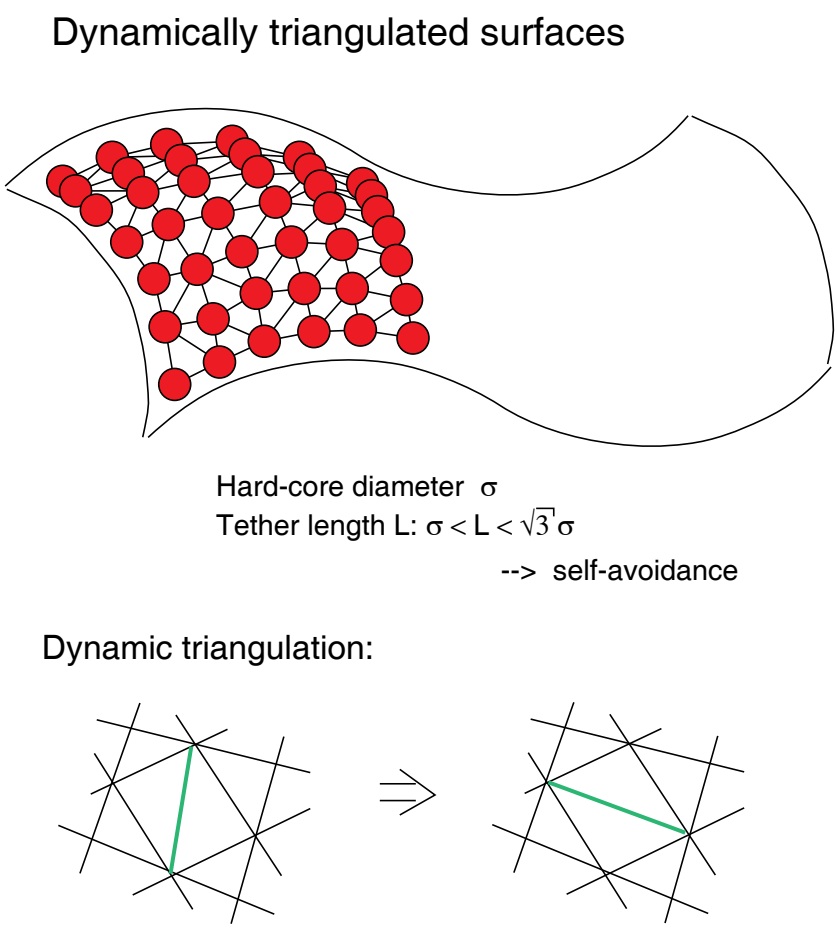


Fig. 12: *A triangulated network model of a fluctuating surface. Top: Hard spheres connected by bonds of maximum extension ℓ are used to describe self-avoiding membranes. Bottom: The Monte Carlo step, which makes the triangulation dynamic, is required to model fluid membranes.*

- The instability argument for the position of the phase transitions, as well as the other calculations which lead to the phase diagram of Fig. 9 apply for $\kappa \gg k_B T$. However, microemulsions and sponge phases typically occur in systems with small bending rigidities of $\kappa \simeq k_B T$.

- Sponge phases with random geometry and bicontinuous *cubic* phases with minimal surface structure cannot be distinguished in the "Fluctuating Membranes" approach. Therefore, few predictions are available from this approach for the structure of the bicontinuous phases.

Simulation method — The simulation of highly dilute sponge phases, with characteristic domain sizes on the order of 10-100nm is impossible on the basis of a molecular model, since it would require an enormous number of "solvent molecules". Therefore, the most appropriate model is again the random surface model discussed above. In order to make this model suitable for simulations, the continuous surface has to be approximated by a network of vertices and bonds, see Fig. 12. A triangular network is usually used because it provides the most homogeneous and isotropic discretization of the surface. A Monte Carlo step then consists of a

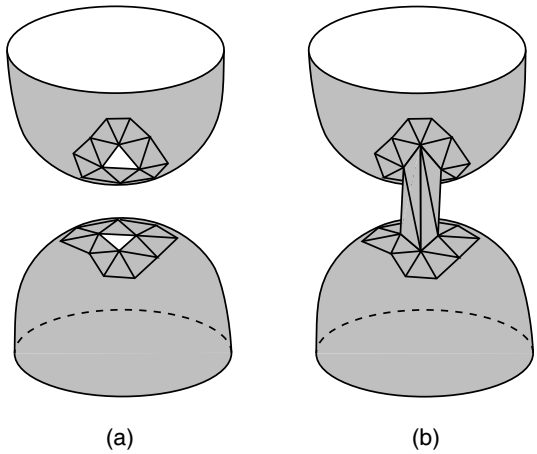


Fig. 13: *Schematic representation of the Monte Carlo step, which is used to change the topology of a randomly triangulated surface. For details see text.*

random displacement of a randomly selected vertex. This step is accepted with the probability determined by the Boltzmann weight, as long as the vertices remain within the maximum bond lengths with their neighbors. The energy, which appears in the Boltzmann weight is the curvature energy, which can be discretized in the form [50, 51]

$$E_b = \lambda_b \sum_{<ij>} (1 - \mathbf{n}_i \cdot \mathbf{n}_j) \tag{38}$$

where $\mathbf{n}_i$ and $\mathbf{n}_j$ are the normal vectors of neighboring triangles, and the sum runs over all pairs of neighboring triangles. The coupling constant λ_b in Eq. (38) is related to the bending rigidity and saddle-splay modulus by $\kappa = \sqrt{3}\lambda_b/2$ and $\bar{\kappa} = -4\kappa/3$ [52]. Other discretizations are discussed in Ref. [53].

Such a model has first been suggested and simulated for *polymerized membranes* by Kantor & Nelson (1987) [50, 51]. When hard spheres of diameter σ_0 are placed on the vertices, and the bond lengths ℓ are restricted to be $\ell \leq \sqrt{3}\sigma_0$, the surface is *self-avoiding*, since an arbitrary sphere does *not* fit through the holes of the network, so that no interpenetration of different parts of the network is possible.

For a study of fluid membranes, the connectivity of the membrane cannot remain fixed during the simulation, because otherwise a diffusion of vertices within the membrane is not possible. Therefore, dynamically triangulated surfaces [11, 54] have to be used in this case. The essential step of the dynamic triangulation procedure is shown at the bottom of Fig. 12. Among the four vertices of two neighboring triangles, the "diagonal" bond is switched from one of the two possible positions to the other. This bond-switching is only allowed if the vertices remain connected to at least three neighbors after the switch. Also, the distance between the newly connected vertices has to be smaller than the maximum bond length. This Monte Carlo step has the advantages that

- is is local, i.e. only the vertices of two neighboring triangles are involved, and
- it guarantees that the network retains its two-dimensional connectivity during the whole simulation run.

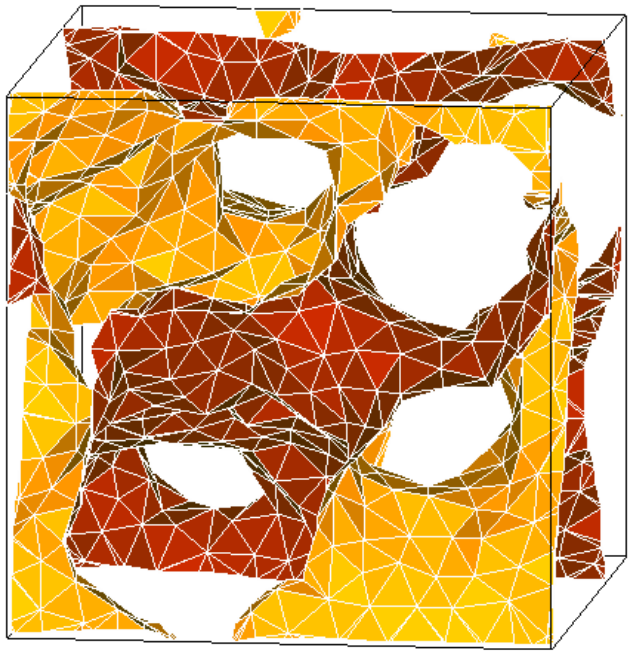


Fig. 14: *A typical conformation of the membrane in a sponge phase, as obtained from Monte Carlo simulations of randomly triangulated surfaces (Gompper & Kroll 1998) [55]. The two sides of the membrane are colored differently to emphasize the bicontinuous structure.*

Finally, in order to study membrane ensembles with fluctuating topology, a Monte Carlo step is required which changes the topology. Such a step is shown schematically in Fig. 13. It consists of removing two triangles from the network, which are located sufficiently closely in the embedding space that a prism of six new triangles can be inserted without exceeding the maximum bond lengths to form a passage-like connection between the previously disjoint membrane patches [55, 54]. Obviously, the inverse step is also possible, where a passage is removed and two triangles are inserted to close the holes in the network.

Simulation Results — A typical conformation of the membrane in the region of the parameter space, where the sponge phase is stable, is shown in Fig. 14. This configuration shows all the characteristic features of a sponge phase, which have been discussed above, such as the bicontinuous structure and the local saddle-like shape of the membrane.

A more quantitative analysis is possible by calculating thermal averages for many different values of the parameters, and then constructing a phase diagram, as shown in Fig. 15. The Monte Carlo data [55] for the phase boundaries are found to follow the logarithmic dependence expected from Eq. (16). Furthermore, the slopes of these phase boundaries in a semi-logarithmic plot are quite consistent with the $\bar{\alpha} = 10/3$, while $\bar{\alpha} = 0$ can clearly be ruled out. Thus, the Monte Carlo results strongly support the "Fluctuating Membranes" predictions, with the value of $\bar{\alpha}$ predicted in Refs. [33, 34].

Another quantity, which can be obtained from the simulations and compared directly with the analytical predictions is the osmotic pressure p . In the analytical approaches, it is obtained

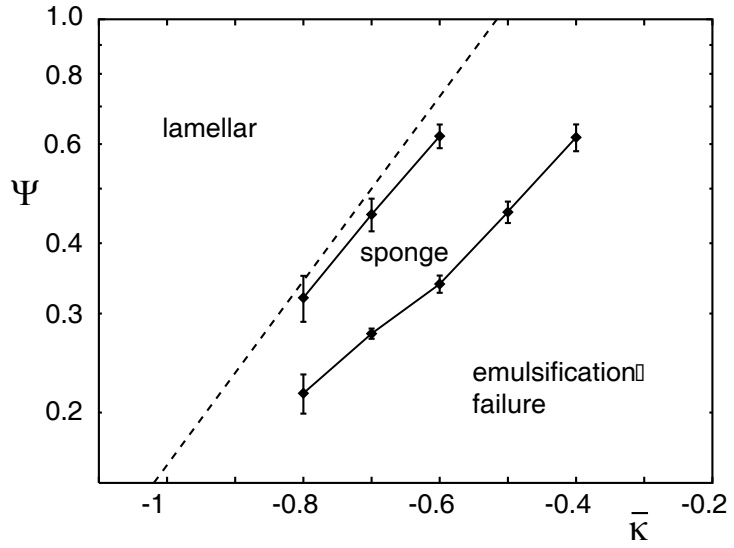


Fig. 15: *Phase diagram of random surfaces controlled by curvature elasticity, as obtained from Monte Carlo simulations of randomly triangulated surfaces with $\kappa \simeq 1.7 k_B T$ (Gompper & Kroll 1998) [55]. The dashed line shows the theoretical prediction (16), with $\bar{\alpha} = 10/3$.*

from the free-energy density f by differentiation,

$$\begin{aligned} \frac{p}{k_B T} &= \Psi \frac{\partial f}{\partial \Psi} - f \\ &= [A'' + B'' \ln(\Psi)] \Psi^3 \qquad \text{(fluctuating membranes)} \end{aligned} \tag{39}$$

and in the case of the free-energy density (36) has the same functional form as f itself. The simulation results (Gompper & Kroll 1998) [55] are compared with this prediction in Fig. 16. Again, the simulation results are in very good agreement with the "Fluctuating Membranes" prediction.

Bulk scattering curves for three different membrane volume fractions Ψ are shown in Fig. 17. The peak position k_0 moves out and the peak broadens, roughly linearly with the membrane volume fractions Ψ, as expected from the variational GRF model, compare Eq. (35). However, the dimensionless length-scale rate $k_0 \xi$ is *not* independent of Ψ, in contrast to the GRF prediction.

5 Experimental Results

Phase behavior of binary systems — The elastic properties of an amphiphilic bilayer can be rather easily related to the elastic properties of the two monolayers of which the bilayer is composed (Petrov et al. 1978, Porte 1989) [57, 58]. This is done by writing the bilayer energy of a spherical cap as

$$\begin{aligned} E_b = \int dS \quad & \left\{ \left[2\kappa \left(\frac{1}{R+\epsilon} + c_0 \right)^2 + \bar{\kappa} \frac{1}{(R+\epsilon)^2} \right] \frac{(R+\epsilon)^2}{R^2} \right. \\ & \left. + \left[2\kappa \left(\frac{1}{R-\epsilon} - c_0 \right)^2 + \bar{\kappa} \frac{1}{(R-\epsilon)^2} \right] \frac{(R-\epsilon)^2}{R^2} \right\} \end{aligned} \tag{40}$$

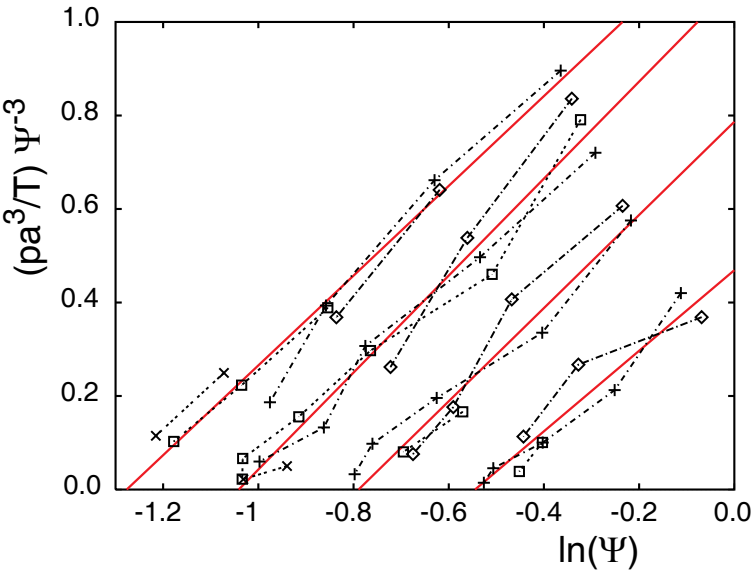


Fig. 16: *Scaled osmotic pressure* $(pa^3/k_BT)\Psi^{-3}$ *of random surfaces controlled by curvature elasticity, as obtained from Monte Carlo simulations of randomly triangulated surfaces for* $\kappa \simeq 1.0k_BT$ *(Gompper & Kroll 1998) [55]. Here,* a *is the membrane thickness. Data sets from right to left correspond to* $\bar{\kappa} = -0.4k_BT$, $\bar{\kappa} = -0.5k_BT$, $\bar{\kappa} = -0.6k_BT$ *and* $\bar{\kappa} = -0.7k_BT$. *The solid red lines indicate the theoretical prediction (36).*

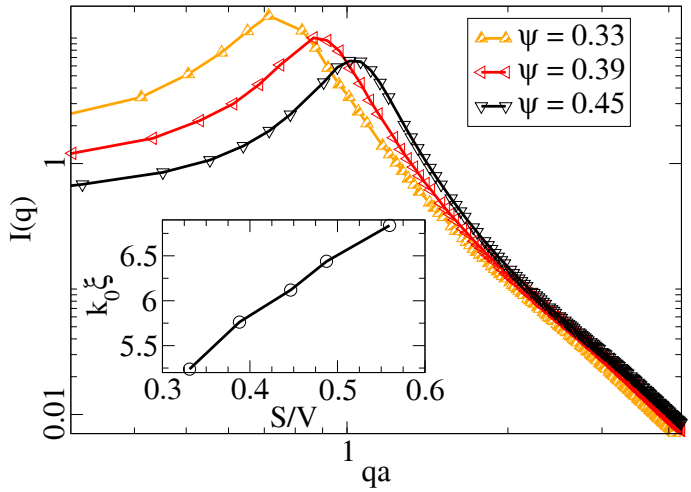


Fig. 17: *Bulk scattering intensity for three membrane volume fractions* $\Psi = aS/V$ *(see legend), as obtained Monte Carlo simulations for* $\kappa \simeq 3.5k_BT$ *and* $\bar{\kappa} = -0.6k_BT$. *The inset shows the length-scale ratio* $k_0\xi$ *as a function of* S/V. *From Peltomäki et al. (2012) [56].*

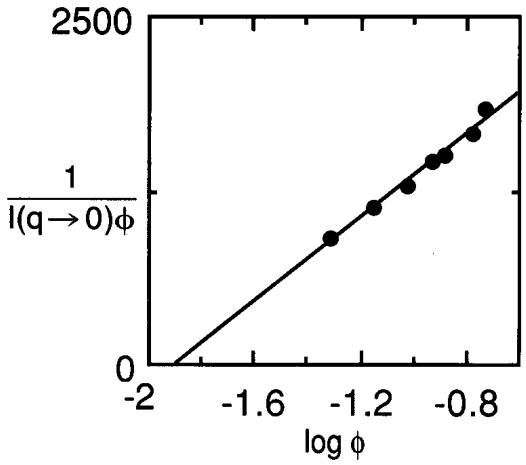


Fig. 18: *Experimental results for the film scattering intensity $I(q)$ in the limit of zero wave vector* $\mathbf{q}$ *in the system AOT–water–$NaCl$ (Skouri et al. 1991) [60].*

where R is the radius of curvature of the bilayer midsurface, ϵ is the thickness of a monolayer, and the integral extends over the surface of the midplane. The expansion of this expression in powers of $1/R$, together with a similar expansion of the energy of a cylindrical or saddle-shaped surface, gives again the form of the curvature energy, but now with

$$c_0^{(bi)} = 0 \quad , \quad \kappa^{(bi)} = 2\kappa \quad , \quad \bar{\kappa}^{(bi)} = 2\bar{\kappa} + 8\kappa c_0 \epsilon \tag{41}$$

Thus, the spontaneous curvature of a bilayer vanishes due to symmetry, but its saddle-splay modulus is related to the spontaneous curvature of the monolayers.
Now, the spontaneous curvature of monolayer has been shown experimentally (Strey 1994) [59] in ternary microemulsions to follow a linear temperature dependence

$$c_0^{(mono)} \sim (T - \bar{T}) \tag{42}$$

over a wide range of temperatures, where $\bar{T}$ is the hydrophobic-hydrophilic-balance temperature, which corresponds to balanced microemulsions. Therefore, $\bar{\kappa}^{(bi)} \sim (T - T^*)$, where T^* is a constant, and we expect the lamellar-to-sponge and the emulsification failure phase transitions to follow the relation

$$\ln(\Psi) \sim (T - T^*) \tag{43}$$

which corresponds to a straight line in a semi-logarithmic representation of the phase diagram. The experimental phase diagram is indeed consistent with such a behavior, as can be seen from Fig. 3.

Small-Angle Scattering — In the limit of very small wave vectors q, the scattering intensity in *film contrast* is determined by the osmotic compressibility, so that [32]

$$\begin{aligned} I(q \to 0) &= \Psi \left(\frac{\partial p}{\partial \Psi} \right)^{-1} \\ &\sim \Psi^{-1} [A''' + B''' \ln(\Psi)]^{-1} \quad \text{(fluctuating membranes)} \end{aligned} \tag{44}$$

This expression is compared to experimental data [32, 60] for the system AOT–water–$NaCl$ in Fig. 18. The data nicely agree with the behavior predicted behavior for "Fluctuating Membranes".

6 Beyond "Simple" Amphiphilic Systems

6.1 Amphiphilic Block Copolymers in Ternary Mixtures

Fig. 19: *Effect of small amounts of amphiphilic block copolymer PEP-PEO (compare Fig. 21 below) on the phase behavior of balanced microemulsions. Far left: oil-water coexistence. Left: 7 weight % surfactant $C_{10}E_4$ produce a three-phase coexistence with a microemulsion phase in the center. Right: 0.5 weight % of PEP*5*-PEO*5 *increase the microemulsion volume two-fold. Far right: 1.0 weight % of PEP*5*-PEO*5 *increases the microemulsion volume three-fold (preparation and photo: J. Allgaier).*

It has been found very recently that the addition of small amounts of amphiphilic block copolymers to balanced microemulsions has a dramatic effect on the phase behavior (Jakobs et al. 1999) [61]. As shown in Fig. 19, the ability of a short-chain surfactant to solubilize oil and water into a macroscopically homogeneous and isotropic microemulsion phase increases dramatically when a small percentage of the total surfactant weight is added as amphiphilic block copolymer. The block copolymer used in this study is polyethyleneoxide–polyethylenepropylene PEOx-PEPy, where x and y denote the molecular weights of each block in kg/mol. The chemical structure of the polymers mimics the structure of the non-ionic C_iE_j surfactant, and PEP10-PEO10 corresponds roughly to $C_{715}E_{230}$. Note that while the microemulsion is completely transparent without block copolymer, it becomes more and more opaque with increasing block-copolymer concentration. This indicates that the scale of the oil- and water domains is increasing with increasing block-copolymer concentration, and it approaches the wavelength of light at the highest polymer concentrations in Fig. 19.

A typical conformation of a polymer-decorated membrane is shown in Fig. 20 on the mesoscopic scale. The block copolymers are incorporated into the surfactant-monolayer membrane, such that the hydrophilic block is located in the water domains, while the hydrophobic block is in the oil domains.

The dramatic effect of amphiphilic block copolymers on the microemulsion can be understood on the basis of the membrane-curvature model [63, 64, 65]. Polymer chains anchored to a

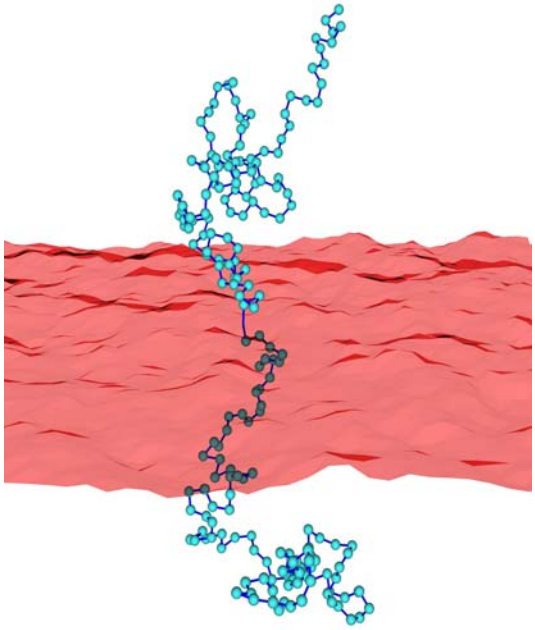


Fig. 20: *Typical conformation of a polymer-decorated membrane (Auth & Gompper 2003) [62]. For details see text.*

membrane modify the curvature elasticity such that (Hiergeist & Lipowsky 1996, Eisenriegler et al. 1996, Auth & Gompper 2003) [66, 67, 62]

$$\kappa_{eff} = \kappa + \frac{k_B T}{12}\left(1+\frac{\pi}{2}\right)\sigma(R_w^2+R_o^2) \tag{45}$$

$$\bar{\kappa}_{eff} = \bar{\kappa} - \frac{k_B T}{6}\sigma(R_w^2+R_o^2) \tag{46}$$

where σ is the grafting density of the polymer, i.e. the number of polymer anchoring points per unit membrane area, and R_w^2 and R_o^2 are the mean-squared end-to-end distances of the polymer blocks in the water- and oil-domains, respectively (compare Fig. 21). Eqs. (45) and (46) apply in the so-called mushroom regime, where the polymer density on the membrane is so small that the polymer coils do not interact directly with each other. Thus, the bending rigidity and the saddle-splay modulus increase *linearly* with the scaled polymer grafting density $\sigma(R_o^2+R_w^2)$ in this regime. It should be noticed that this is a very *small effect*, since even at the overlap concentration, where $\sigma(R_o^2+R_w^2)=2$ for a symmetric block copolymer, the increment of κ and $|\bar{\kappa}|$ is only a fraction of k_BT. The results (45) and (46) can be compared, for example, with the elastic properties of solid elastic sheets, where the bending rigidity scales with the sheet thickness ℓ_0 as $\kappa \sim \ell_0^3$ [68].

The increase of the effective bending rigidity κ_{eff} in Eq. (45) and of the magnitude of $\bar{\kappa}_{eff}$ in Eq. (46) can be understood from the entropy loss of each polymer coil in a confined geometry due to the restriction of the number of accessible configurations. This entropic repulsion suppresses membrane fluctuations *and* disfavors saddle-like conformations near the anchoring points. A simple, heuristic argument for the suppression of saddle-like conformations has been given by Milner & Witten (1988) [69] for polymers in the brush regime, where there is a strong

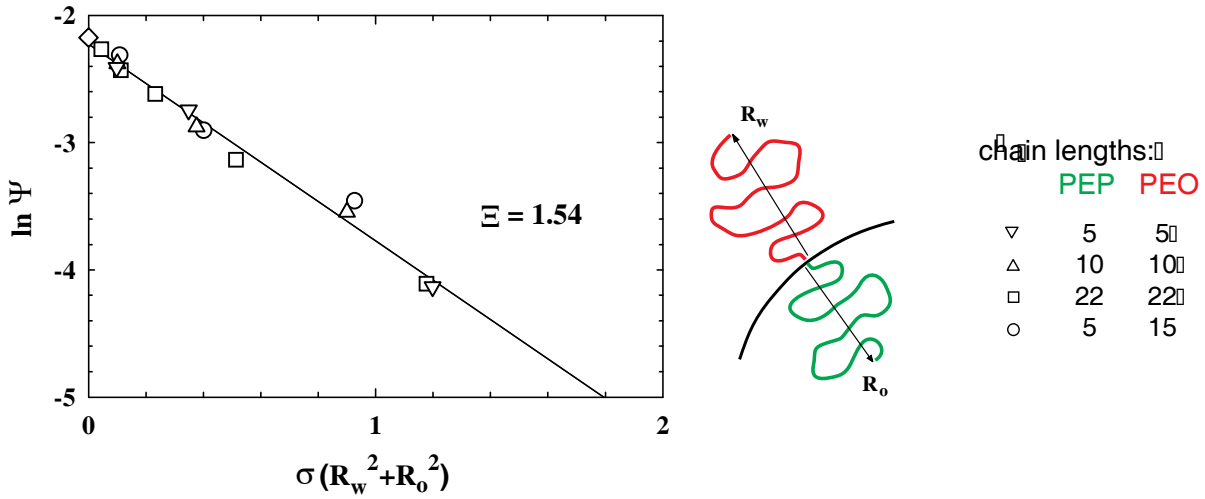


Fig. 21: *Phase diagram of a microemulsion of water–decane–$C_{10}E_4$ with amphiphilic block copolymer. The membrane volume fraction Ψ of the microemulsion at three-phase coexistence is shown as a function of the scaled polymer grafting density on the membrane (Endo et al. 2000) [63]. The different symbols correspond to different molecular weights (in kg/mol) of the four different PEP-PEO block copolymers. Average end-to-end distances of the two blocks in the water and oil subphase are denoted by R_w and R_o, respectively.*

overlap between neighboring chains on the membrane. Consider the volume V of a thin shell of thickness ℓ_0 above a surface of area A bent with local curvatures c_1 and c_2. The ratio $V/(A\ell_0)$ is the ratio of volume available for chains in a layer of thickness ℓ_0 and grafting area A in the bent and unbent geometries, and may be expanded as

$$\frac{V}{A\ell_0} = 1 + H\ell_0 + \frac{1}{3}K\ell_0^2 + ... \tag{47}$$

where H and K are the mean and Gaussian curvatures, respectively. This relation can be checked easily for a spherical and a cylindrical surface. Then for a saddle surface, $H = 0$ and $K = -1/R^2$, so that $V/(A\ell_0) \simeq 1 - \frac{1}{3}(\ell_0/R)^2$; the bent thin shells have therefore *less space* available for monomers than the unbent shells, and the polymer chains must stretch upon bending, which costs free energy. The same argument should also apply to the mushroom regime.

Thus, from the combination of Eqs. (16) and (46), the prediction

$$\begin{aligned} \ln(\Psi/\Psi^*) &= \frac{4\pi}{\bar{\alpha}}\frac{1}{6}\sigma(R_w^2 + R_o^2) \\ &= -\frac{\pi}{5}\sigma(R_w^2 + R_o^2) \end{aligned} \tag{48}$$

for the dependence of the emulsification boundary on the scaled polymer density is obtained. This result is compared with experimental data for a number of different block-copolymer lengths in Fig. 21. The exponential dependence of the membrane volume fraction Ψ at three-phase coexistence is very nicely confirmed. In fact, this *exponential* dependence is responsible that the small increment in the magnitude of the saddle-splay modulus has the dramatic effects on the phase behavior shown in Fig. 19. Also, the data for different polymer lengths scale exactly as predicted by Eq. (48). A fit of the experimental data in Fig. 21 to the functional form

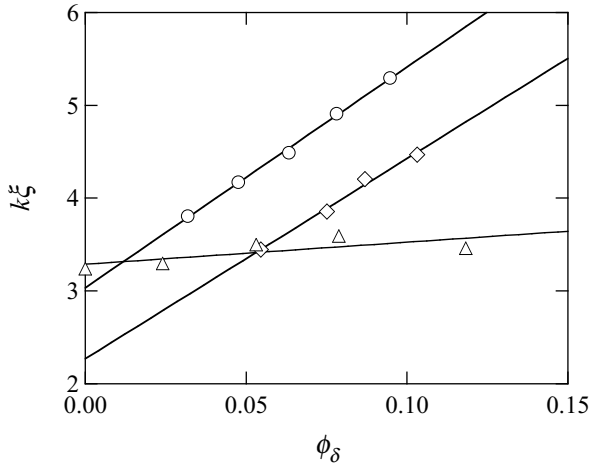


Fig. 22: *Dimensionless ratio $k_0\xi$ of the two characteristic length scales of a microemulsion, as a function of the polymer content, ϕ_δ, in the mixture of both amphiphiles, which is proportional to the polymer grafting density on the membrane (Gompper et al. 2001) [70]. The two data sets marked by circles ($\circ$) and diamonds ($\diamond$) were obtained for PEP5-PEO5 at fixed membrane volume fractions of $\Psi = 0.12$ and $\Psi = 0.08$, respectively. The data set marked by triangles ($\triangle$) were taken for PEP10-PEO10 along three-phase coexistence.*

of Eq. (48) yields a prefactor $\Xi = 1.54$, which is a bit more than a factor two larger than the theoretical result $\pi/5 \simeq 0.628$. A possible reason for this discrepancy is the theoretical result (45) has been obtained for ideal chains, while the polymer chains are self-avoiding in the real system. This is one of the few examples, where a *quantitative* comparison of experimental data and theoretical results has been achieved for these complex, disordered, and mesoscopically structured fluid phases.

It is possible to go one step further and also investigate the scattering behavior [70]. In this case, the Monte Carlo simulations of triangulated surfaces predicts (Peltomäki et al. 2012) [56],

$$k_0\xi = 7.39(0.15\kappa^R + 0.85\bar{\kappa}^R) \sim \sigma(R_o^2 + R_w^2) \tag{49}$$

where the later relation follows from Eqs. (45) and (46). The experimental data for this characteristic number are shown in Fig. 22. The data for fixed surfactant concentration, i.e. for fixed membrane area, follow precisely the behavior of Eq. (49). It is important to notice, however, that although the lines have equal slope, as expected from Eq. (49), they are shifted relative to each other. The data along the coexistence line still depend linearly on the polymer content ϕ_δ in the mixture of both amphiphiles, but even the slope is different. This is due to the logarithmic renormalization of the bending rigidity [70, 56].

The extraction of bending rigidities has been applied recently to other microemulsion systems, in particular to microemulsions containing supercritical CO_2 instead of oil [71, 72].

6.2 Interfaces and Walls

So far, we have focused on bulk phases. However, microemulsions can of course coexist with other phases, such as a water-rich and an oil-rich phase, see Sec. 2. Particularly important is also the behavior near wall, because microemulsions are used, for example, for cleaning processes.

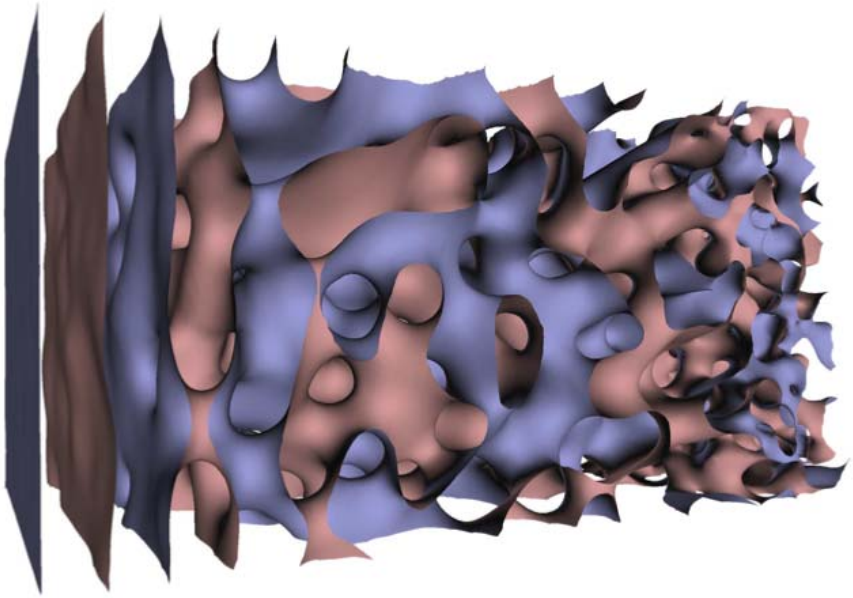


Fig. 23: *Microemulsion near a planar wall (left). The wall induces a lamellar structure close to it (Kerscher et al. 2011) [73].*

Results of Monte Carlo simulations employing the Ginzburg-Landau model described in Sec. 3.3 for the microemulsion structure near a wall shown in Fig. 23. The planar wall induces a few layers of the lamellar phase near the wall. The scattering intensity for a small angle neutron scattering study under grazing incidence (GISANS) , which is predicted from this structure, agrees very well the recent experimental results (Kerscher et al. 2011) [73]. An interesting theoretical prediction is that the "transition" from perfect lamellae to the bicontinuous structure occurs through a layer of perforated lamellae, where neighboring membranes are connected by catenoid-like passages.
Another interesting interface occurs between different domains of the same phase of ordered bicontinuous structure, which are well described by triply periodic minimal surfaces, compare Sec. 3.2. As for any other crystalline material, different kind of interfaces can be distinguished, in particular tilt and twist grain boundaries, where the two grains are rotated with respect to each other with a rotation axis parallel and perpendicular to the interface, respectively. Results of a free-energy minimization in the Ginzburg-Landau model are show in Fig. 24. The interface is difficult to see, because the membrane structure adapts to the constraints very well. This makes it plausible that also the interface energy (corresponding to the surface tension between fluid interfaces) is very small, which is indeed the result of the calculation [74]. This implies that grain boundaries should be easily excited by external forces.

7 Summary and Conclusions

Research in the field of Soft Matter systems is an interdisciplinary endeavor. The work on random surfaces and fluctuating membranes brings together mathematicians (minimal surfaces), theoretical physicists (statistical physics, field theory), experimental physicists (neutron scat-

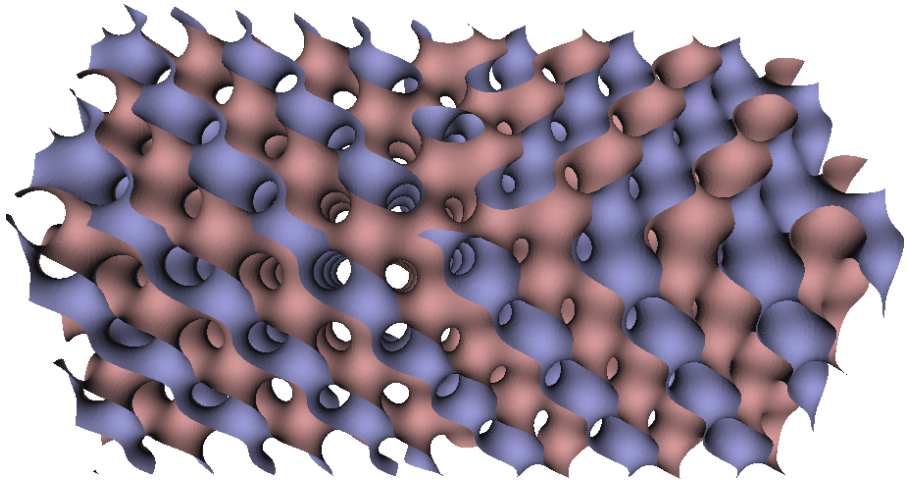


Fig. 24: *Interface between two different domains of the gyroid phase. The interface, a twist grain boundary, is vertical, with the domains of the TPMS to the left and right, rotated by an angle of* 90^o *with respect to each other (Belushkin & Gompper 2009) [74].*

tering, microscopy), chemists (synthesis, phase behavior), biologists (cell membranes), and material scientists (nanoscale materials).

Over the last twenty years, the joint effort of researches from these different disciplines has lead to significant progress in our understanding of amphiphilic systems. For the "simple" system of an ensemble of membranes without internal structure, the qualitative behavior is reasonably well understood. On the quantitative level, there remains still a lot of work to be done. As the discussion of the previous sections has demonstrated, the results of different approximate approaches do not always agree. Also, experiments are often done in rather complex systems, where several different physical mechanisms are at work simultaneously, which make the interpretation of the results difficult. Therefore, even the "simple" system will require much more work before a satisfactory level of understanding has been reached.

The comparison of theoretical models, computer simulation results, and experimental data indicates that the "fluctuating interfaces" approach works very well. The effect of the logarithmic renormalization of the elastic curvature moduli now seems to be well established. Simulations show that also scattering intensities of microemulsion or sponge phases can be predicted from this approach. The results are predicted surprisingly well by the "Gaussian random field" approach, but show important qualitative and quantitative differences.

In the recent past, and undoubtedly in the future, the trend will go towards more complex systems. With the increasing number of components, the system will become more flexible, with properties which can be controlled and manipulated externally. Here, the contact with biology will certainly intensify. For Statistical Physics, it will be a very interesting task to predict the properties of such complex systems quantitatively.

Appendices

A Correlation Functions of Gaussian Free-Energy Functionals

The partition function Z for any free-energy functional $\mathcal{F}(\Phi)$ is a functional integral over all order parameter configurations $\Phi(\mathbf{r})$. This integral is written formally as

$$Z_0 = \int \mathcal{D}\Phi \exp[-\mathcal{F}] \tag{50}$$

By adding auxiliary fields $H(\mathbf{r}$ to the free-energy functional, correlation functions can be derived from the (modified) partition function

$$Z\{H\} = \int \mathcal{D}\Phi \exp\left[-\mathcal{F} + \int d^3r H(\mathbf{r})\Phi(\mathbf{r})\right] \tag{51}$$

by functional differentiation. For example, the two-point order parameter correlation function is obtained as

$$G(\mathbf{r}-\mathbf{r}') =< \Phi(\mathbf{r})\Phi(\mathbf{r}') >= \frac{1}{Z_0}\frac{\delta^2 Z\{H\}}{\delta H(\mathbf{r})\delta H(\mathbf{r}')}|_{H=0}\,. \tag{52}$$

The exact evaluation of the functional integral (51) is, however, only possible in a few special cases; in particular, this is possible for free-energy functionals of the quadratic form

$$\mathcal{F}_0 = \int d^3r \int d^3r'\, w(|\mathbf{r}-\mathbf{r}'|)\Phi(\mathbf{r})\Phi(\mathbf{r}') \tag{53}$$

$$= \int d^3q\, w(|\mathbf{q}|)\Phi(\mathbf{q})\Phi(-\mathbf{q}) \tag{54}$$

For the free-energy functional of Eq. (19), a Fourier transformation yields

$$w(q) = \frac{1}{2}\left(c_0 q^4 + b_0 q^2 + a_0\right) \tag{55}$$

In all other cases of non-Gaussian functionals, one has to resort to a perturbative approach [75, 44].

We want to focus here on the calculation of correlation functions. For the functional (53), the partition function (51) is given by

$$Z\{H\} = \int \mathcal{D}\Phi\, \exp\left[\int d^3q\,[-w(|\mathbf{q}|)\Phi(\mathbf{q})\Phi(-\mathbf{q}) + H(\mathbf{q})\Phi(-\mathbf{q})]\right] \tag{56}$$

By completing the square in the exponent, we obtain

$$Z\{H\} = \int \mathcal{D}\Phi\, \exp\left[\int d^3q \left[-w(q)\left|\left(\Phi(\mathbf{q}) - \frac{H(\mathbf{q})}{2w(q)}\right)\right|^2 - \frac{H(\mathbf{q})H(-\mathbf{q})}{4w(q)}\right]\right] \tag{57}$$

Here, we have used that the order parameter $\Phi(\mathbf{r})$ and the auxiliary fields $H(\mathbf{r})$ are real numbers, so that $\Phi(-\mathbf{q}) = \Phi(\mathbf{q})^*$ and $H(-\mathbf{q}) = H(\mathbf{q})^*$. By introducing the new integration variables $\tilde{\Phi}(\mathbf{q}) = \Phi(\mathbf{q}) - H(\mathbf{q})/2w(q)$, we obtain from Eq. (57)

$$Z\{H\} = \exp\left[-\frac{H(\mathbf{q})H(-\mathbf{q})}{4w(q)}\right] \int \mathcal{D}\tilde{\Phi}\, \exp\left[-\int d^3q\, w(q)\tilde{\Phi}(\mathbf{q})\tilde{\Phi}(-\mathbf{q})\right] \tag{58}$$

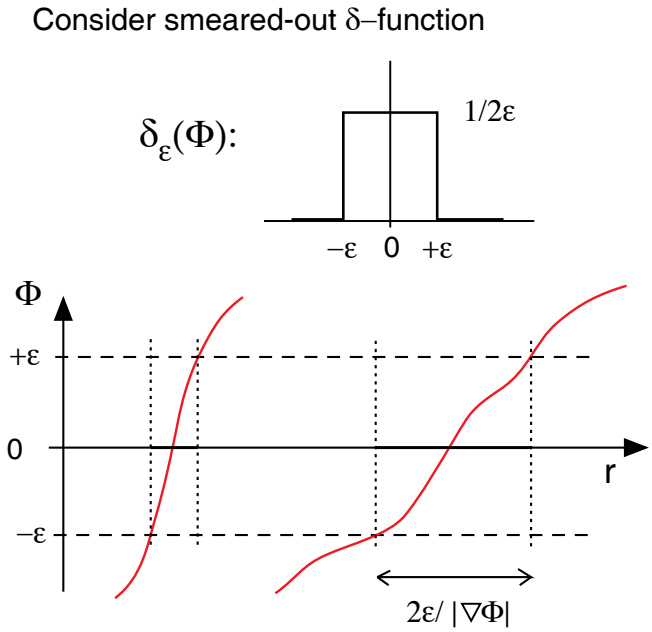


Fig. 25: *In order to calculate the surface area of level surfaces of Gaussian random fields, the δ-function in Φ has to be corrected for the different local gradients of the order-parameter field. For details see text.*

The remaining functional integral is the partition function Z_0, compare Eq. (50), which is *independent* of the fields $H(\mathbf{q})$, and therefore is not required for the calculation of correlation functions. The evaluation of the functional derivatives in Eq. (52) is now straightforward and yields

$$G(\mathbf{q}) = \frac{1}{2w(q)} \tag{59}$$

Together with Eq. (55), this implies Eq. (20).

B Calculation of the Surface Density S/V for Gaussian Random Fields

We briefly sketch the calculation of the surface density S/V for Gaussian random. It is clear that the surface density must be related to the average of $\delta(\Phi(\mathbf{r})-\alpha)$. However, note that in contrast to the particle density of molecules or colloidal particles, which is given by $\rho(\mathbf{r}) = \sum_i \delta(\mathbf{r}-\mathbf{r}_i)$, the δ-function for the level surfaces acts on the Φ-variable, rather than the space variable $\mathbf{r}$ is in the usual case. Therefore, we need the mathematical identity

$$\delta(\Phi(\mathbf{r}) - \alpha) = \frac{1}{|\nabla\Phi(\mathbf{r})|}\delta(\mathbf{r} - \mathbf{r}_0) \tag{60}$$

where $\mathbf{r}_0$ is the location of the level surface. The surface density is therefore given by

$$S/V = \langle |\nabla\Phi(\mathbf{r})|\, \delta(\Phi(\mathbf{r}) - \alpha)\rangle \tag{61}$$

There is another, more intuitive way to arrive at the same conclusion. Let us consider a smeared-out δ-function of width ϵ, as shown in Fig. 25 (upper part). The contribution to the average of the function δ_ϵ now clearly depends on the gradient of Φ; if the gradient is small (large), the contribution is large (small), as shown in Fig. 25 (lower part). Since for the calculation of S/V the contribution of all parts of the surface must be the same — independent of the local gradient of Φ — this effect has to be divided out, which gives again Eq. (61).
In order to calculate the average (61), we need the joint probability distribution function

$$P[\Phi(\mathbf{r}) = \alpha, \nabla\Phi(\mathbf{r}) = \mathbf{v}] = \langle \delta(\Phi(\mathbf{r}) - \alpha)\, \delta(\nabla\Phi(\mathbf{r}) - \mathbf{v}) \rangle \tag{62}$$

It is shown in detail in Appendix C that

$$P[\alpha, \mathbf{v}] = P_0 \exp\left[-\frac{\alpha^2}{2} - \frac{v^2}{2\sigma_v^2} \right] \tag{63}$$

with

$$\sigma_v = \frac{1}{3} \int \frac{d^3q}{(2\pi)^3}\, q^2 \nu(q) \equiv \frac{1}{3} \langle q^2 \rangle_\nu \tag{64}$$

and normalization

$$P_0 = (2\pi)^{-2} \sigma_v^{-3} \tag{65}$$

This implies

$$\begin{aligned} S/V &= \int d^3v |v| P(\alpha, \mathbf{v}) \\ &= 4\pi P_0 \int_0^\infty dv v^3 \exp[-\alpha^2/2 - v^2/(2\sigma_v^2)] \\ &= \frac{2}{\pi} \exp[-\alpha^2/2] \sqrt{\frac{1}{3}\langle q^2 \rangle_\nu} \end{aligned} \tag{66}$$

This is just the result (26).

C Joint Distribution of Gaussian Random Fields

Joint probability distribution functions are required for the calculation of geometrical averages of level surfaces of Gaussian random fields, see Appendix B. First, the probability distribution

$$P[\Phi(\mathbf{r}) = \alpha, \nabla\Phi(\mathbf{r}) = \mathbf{v}] = \langle \delta(\Phi(\mathbf{r}) - \alpha)\, \delta(\nabla\Phi(\mathbf{r}) - \mathbf{v}) \rangle \tag{67}$$

has to be calculated [39, 6]. Since this probability distribution only contains a single spatial variable, and the system is homogeneous on average, $\mathbf{r} = 0$ can be chosen without loss of generality. The representation

$$\delta(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp(ikx) \tag{68}$$

can now be used to obtain

$$P[\alpha, \mathbf{v}] = \int \frac{d\omega}{2\pi} \int \frac{d^3k}{(2\pi)^3} \left\langle e^{i\omega(\Phi(\mathbf{r}) - \alpha)} e^{i\mathbf{k}\cdot(\nabla\Phi(\mathbf{r}) - \mathbf{v})} \right\rangle \tag{69}$$

With the Fourier-representation

$$\Phi(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3}\Phi(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}} \tag{70}$$

this becomes

$$P[\alpha,\mathbf{v}] = \int \frac{d\omega}{2\pi}\int \frac{d^3k}{(2\pi)^3}e^{-i\omega\alpha - i\mathbf{k}\cdot\mathbf{v}}\left\langle \exp\int \frac{d^3q}{(2\pi)^3}[i\omega\Phi(\mathbf{q}) - \mathbf{k}\cdot\mathbf{q}\Phi(\mathbf{q})]\right\rangle \tag{71}$$

at $\mathbf{r} = 0$. The average in Eq. (71) has to be evaluated with the Boltzmann weight

$$\exp[-\mathcal{H}_0] = \exp\left\{-\frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\nu(q)^{-1}|\Phi(\mathbf{q})|^2\right\} \tag{72}$$

which implies in particular

$$\langle\Phi(\mathbf{q})\Phi(\mathbf{q}')\rangle = \nu(q)(2\pi)^3\delta(\mathbf{q}+\mathbf{q}') \tag{73}$$

For a Gaussian-distributed variable $x \in \mathbb{C}$, the relation

$$\langle e^{ikx}\rangle = \exp(-\frac{1}{2}k^2\langle x^2\rangle) \tag{74}$$

holds exactly. Therefore,

$$\begin{aligned}
&\left\langle \exp\int \frac{d^3q}{(2\pi)^3}[i\omega - \mathbf{k}\cdot\mathbf{q}]\Phi(\mathbf{q})\right\rangle \\
&= \exp\left[\frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\int \frac{d^3q'}{(2\pi)^3}(i\omega - \mathbf{k}\cdot\mathbf{q})(i\omega - \mathbf{k}\cdot\mathbf{q}')\langle\Phi(\mathbf{q})\Phi(\mathbf{q}')\rangle\right] \\
&= \exp\left[-\frac{1}{2}\int \frac{d^3q}{(2\pi)^3}(\omega^2 + (\mathbf{k}\cdot\mathbf{q})^2)\nu(q)\right]
\end{aligned} \tag{75}$$

so that

$$P[\alpha,\mathbf{v}] = \int \frac{d\omega}{2\pi}\int \frac{d^3k}{(2\pi)^3}e^{-i\omega\alpha - i\mathbf{k}\cdot\mathbf{v}}\exp\int \frac{d^3q}{(2\pi)^3}\left[-\frac{1}{2}\omega^2 - \frac{1}{2}(\mathbf{k}\cdot\mathbf{q})^2\right]\nu(q) \tag{76}$$

The remaining integrals are now straightforward. In particular,

$$\int \frac{d^3q}{(2\pi)^3}(\mathbf{k}\cdot\mathbf{q})^2\,\nu(q) = \frac{1}{3}k^2\langle q^2\rangle_\nu \equiv k^2\sigma_v^2 \tag{77}$$

with $\langle ...\rangle_\nu$ defined in Eq. (30), and

$$\int \frac{d^3k}{(2\pi)^3}e^{-i\mathbf{k}\cdot\mathbf{v}}\,e^{-\frac{1}{6}k^2\langle q^2\rangle} = (2\pi)^{-3/2}\sigma_v^{-3}\exp\left[-\frac{1}{2}\frac{v^2}{\sigma_v^2}\right] \tag{78}$$

References

[1] Research in Jülich, Nr. 2/2006, Physics meets Biology.

[2] G. Gompper, J. K. G. Dhont, and D. Richter, Eur. Phys. J. E **26**, 1 (2008).

[3] G. Porte, J. Phys.: Condens. Matter **4**, 8649 (1992).

[4] G. Gompper and M. Schick, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. Lebowitz (Academic Press, London, 1994), vol. 16, pp. 1–176.

[5] W. M. Gelbart, A. Ben-Shaul, and D. Roux, eds., *Micelles, Membranes, Microemulsions, and Monolayers* (Springer-Verlag, Berlin, 1994).

[6] S. A. Safran, *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes* (Addison-Wesley, Reading, MA, 1994).

[7] D. F. Evans and H. Wennerström, *The Colloidal Domain, where Physics, Chemistry, Biology and Technology Meet* (VCH Publishers, New York, 1994).

[8] R. Lipowsky and E. Sackmann, eds., *Structure and dynamics of membranes - from cells to vesicles*, vol. 1 of *Handbook of biological physics* (Elsevier, Amsterdam, 1995).

[9] F. David, P. Ginsparg, and J. Zinn-Justin, eds., *Fluctuating geometries in statistical mechanics and field theory* (Elsevier, Amsterdam, 1996).

[10] U. Seifert, Adv. Phys. **46**, 13 (1997).

[11] G. Gompper and D. M. Kroll, J. Phys.: Condens. Matter **9**, 8795 (1997).

[12] M. E. Cates and M. R. Evans, *Soft and Fragile Matter: Nonequilibrium Dynamics, Metastability and Flow* (Institute of Physics, London, 2000).

[13] J. K. G. Dhont, G. Gompper, and D. Richter, eds., *Soft Matter — Complex Materials on Mesoscopic Scales*, vol. 10 of *Matter and Materials* (Forschungszentrum Jülich, Jülich, 2002).

[14] D. Nelson, T. Piran, and S. Weinberg, eds., *Statistical Mechanics of Membranes and Surfaces* (World Scientific, Singapore, 2004), 2nd ed.

[15] T. A. Witten and P. A. Pincus, *Structured Fluids: Polymers, Colloids, Surfactants* (Oxford University Press, Oxford, 2006).

[16] W. C. K. Poon and D. Andelman, *Soft Condensed Matter Physics in Molecular and Cell Biology* (Taylor & Francis, Boca Raton, 2006).

[17] J. K. G. Dhont, G. Gompper, G. Nägele, D. Richter, and R. G. Winkler, eds., *Soft Matter — From Synthetic to Biological Materials*, vol. 1 of *Key Technologies* (Forschungszentrum Jülich, Jülich, 2008).

[18] J. K. G. Dhont, G. Gompper, P. Lang, D. Richter, M. Ripoll, D. Willbold, and R. Zorn, eds., *Macromolecular Systems in Soft and Living Matter*, vol. 20 of *Key Technologies* (Forschungszentrum Jülich, Jülich, 2011).

[19] R. Strey, R. Schomäcker, D. Roux, F. Nallet, and U. Olsson, J. Chem. Soc. Faraday Trans. **86**, 2253 (1990).

[20] W. Helfrich, Z. Naturforsch. **33a**, 305 (1978).

[21] C. R. Safinya, D. Roux, G. S. Smith, S. K. Sinha, P. Dimon, N. A. Clark, and A. M. Bellocq, Phys. Rev. Lett. **57**, 2718 (1986).

[22] C. R. Safinya, E. B. Sirota, D. Roux, and G. S. Smith, Phys. Rev. Lett. **62**, 1134 (1989).

[23] G. Gompper and D. M. Kroll, Europhys. Lett. **9**, 59 (1989).

[24] R. Strey, W. Jahn, M. Skouri, G. Porte, J. Marignan, and U. Olsson, in *Structure and Dynamics of Strongly Interacting Colloids and Supramolecular Aggregates in Solution*, edited by S.-H. Chen, J. S. Huang, and P. Tartaglia (Kluwer, Dordrecht, 1992), pp. 351–363.

[25] H. T. Davis, J. F. Bodet, L. E. Scriven, and W. G. Miller, in *Physics of Amphiphilic Layers*, edited by J. Meunier, D. Langevin, and N. Boccara (Springer-Verlag, Berlin, 1987), pp. 310–327.

[26] W. Helfrich, Z. Naturforsch. **28c**, 693 (1973).

[27] L. E. Scriven, Nature **263**, 123 (1976).

[28] P.-G. de Gennes and C. Taupin, J. Phys. Chem. **86**, 2294 (1982).

[29] S. A. Safran, D. Roux, M. E. Cates, and D. Andelman, Phys. Rev. Lett. **57**, 491 (1986).

[30] P. B. Canham, J. Theor. Biol. **26**, 61 (1970).

[31] U. S. Schwarz and G. Gompper, Phys. Rev. E **59**, 5528 (1999).

[32] G. Porte, M. Delsanti, I. Billard, M. Skouri, J. Appell, J. Marignan, and F. Debeauvais, J. Phys. II France **1**, 1101 (1991).

[33] L. Peliti and S. Leibler, Phys. Rev. Lett. **54**, 1690 (1985).

[34] W. Cai, T. C. Lubensky, P. Nelson, and T. Powers, J. Phys. II France **4**, 931 (1994).

[35] F. David, in *Statistical Mechanics of Membranes and Surfaces*, edited by D. R. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 1989), pp. 157–223.

[36] D. C. Morse, Phys. Rev. E **50**, R2423 (1994).

[37] G. Gompper and M. Schick, Phys. Rev. Lett. **65**, 1116 (1990).

[38] M. Teubner and R. Strey, J. Chem. Phys. **87**, 3195 (1987).

[39] M. Teubner, Europhys. Lett. **14**, 403 (1991).

[40] N. F. Berk, Phys. Rev. Lett. **58**, 2718 (1987).

[41] N. F. Berk, Phys. Rev. A **44**, 5069 (1991).

[42] P. Pieruschka and S. A. Safran, Europhys. Lett. **22**, 625 (1993).

[43] R. P. Feynman, *Statistical Mechanics* (Addison-Wesley, Reading, MA, 1972).

[44] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena* (Clarendon Press, Oxford, 1992).

[45] P. Pieruschka and S. A. Safran, Europhys. Lett. **31**, 207 (1995).

[46] D. Roux, M. E. Cates, U. Olsson, R. C. Ball, F. Nallet, and A. M. Bellocq, Europhys. Lett. **11**, 229 (1990).

[47] D. Roux, C. Coulon, and M. E. Cates, J. Phys. Chem. **96**, 4174 (1992).

[48] W. Helfrich, J. Phys. France **46**, 1263 (1985).

[49] W. Helfrich, Eur. Phys. J. B **1**, 481 (1998).

[50] Y. Kantor and D. R. Nelson, Phys. Rev. Lett. **58**, 2774 (1987).

[51] Y. Kantor and D. R. Nelson, Phys. Rev. A **36**, 4020 (1987).

[52] J. Lidmar, L. Mirny, and D. R. Nelson, Phys. Rev. E **68**, 051910 (2003).

[53] G. Gompper and D. M. Kroll, J. Phys. I France **6**, 1305 (1996).

[54] G. Gompper and D. M. Kroll, in *Statistical Mechanics of Membranes and Surfaces*, edited by D. R. Nelson, T. Piran, and S. Weinberg (World Scientific, Singapore, 2004), pp. 359–426, 2nd ed.

[55] G. Gompper and D. M. Kroll, Phys. Rev. Lett. **81**, 2284 (1998).

[56] M. Peltomäki, G. Gompper, and D. M. Kroll, preprint (2012).

[57] A. G. Petrov, M. D. Mitov, and A. Derzhanski, Phys. Lett. A **65**, 374 (1978).

[58] G. Porte, J. Appell, P. Bassereau, and J. Marignan, J. Phys. France **50**, 1335 (1989).

[59] R. Strey, Colloid Polym. Sci. **272**, 1005 (1994).

[60] M. Skouri, J. Marignan, J. Appell, and G. Porte, J. Phys. II France **1**, 1121 (1991).

[61] B. Jakobs, T. Sottmann, R. Strey, J. Allgaier, L. Willner, and D. Richter, Langmuir **15**, 6707 (1999).

[62] T. Auth and G. Gompper, Phys. Rev. E **68**, 051801 (2003).

[63] H. Endo, J. Allgaier, G. Gompper, B. Jakobs, M. Monkenbusch, D. Richter, T. Sottmann, and R. Strey, Phys. Rev. Lett. **85**, 102 (2000).

[64] H. Endo, M. Mihailescu, M. Monkenbusch, J. Allgaier, G. Gompper, D. Richter, B. Jakobs, T. Sottmann, R. Strey, and I. Grillo, J. Chem. Phys. **115**, 580 (2001).

[65] G. Gompper, D. Richter, and R. Strey, J. Phys.: Condens. Matter **13**, 9055 (2001).

[66] C. Hiergeist and R. Lipowsky, J. Phys. II France **6**, 1465 (1996).

[67] E. Eisenriegler, A. Hanke, and S. Dietrich, Phys. Rev. E **54**, 1134 (1996).

[68] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Addison-Wesley, Reading, MA, 1959).

[69] S. T. Milner and T. A. Witten, J. Phys. France **49**, 1951 (1988).

[70] G. Gompper, H. Endo, M. Mihailescu, J. Allgaier, M. Monkenbusch, D. Richter, B. Jakobs, T. Sottmann, and R. Strey, Europhys. Lett. **56**, 683 (2001).

[71] M. Klostermann, T. Foster, R. Schweins, P. Lindner, O. Glatter, R. Strey, and T. Sottmann, Phys. Chem. Chem. Phys. **13**, 20289 (2011).

[72] O. Holderer, M. Klostermann, M. Monkenbusch, R. Schweins, P. Lindner, R. Strey, D. Richter, and T. Sottmann, Phys. Chem. Chem. Phys. **13**, 3022 (2011).

[73] M. Kerscher, P. Busch, S. Mattauch, H. Frielinghaus, D. Richter, M. Belushkin, and G. Gompper, Phys. Rev. E **83**, 030401(R) (2011).

[74] M. Belushkin and G. Gompper, J. Chem. Phys. **130**, 134712 (2009).

[75] D. J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena* (World Scientific, Singapore, 1984).

B 3 Diffusion [1]

Jan K.G. Dhont
Institut of Complex Systems ICS-3
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Diffusion relates to the displacement of molecules due to their thermal motion. Diffusion is a general phenomenon that occurs in gases, fluids and solids. Clearly, the interactions between molecules affect their thermally induced displacements. In a crystalline solid, for example, a relatively small diffusing molecule is surrounded by large molecules that reside on average at their crystal lattice sites. These surrounding molecules form a "cage" within which the given small molecule moves around. Depending on the height of the energy barrier set by the interactions with the molecules forming the cage, the tagged molecule occasionally escapes from a cage and moves to a neigbouring cage. In a gas the displacement of a molecule is not hindered by a structured cage of neighbouring molecules, as in a solid. Here, occasional collisions with other molecules in this very dilute system will change the magnitude and direction of the velocity of the tagged molecule. In a fluid the diffusion mechanism is in between that of a solid and a gas: there is a "blurry cage" around each molecule, but this cage is highly dynamic itself, which enhances the cage-escape frequency.
The above described thermal motion of a single molecule in a macroscopically homogeneous system (as depicted in Fig.1a) is referred to as **self diffusion**, where "self" refers to the fact that the dynamics of a single molecule is considered. The diffusive motion of a single molecule can be quantified as follows. Consider a molecule with a prescribed position of its center-of-mass $\mathbf{r}_0$ at time $t = 0$. Let $\mathbf{r}(t)$ denote the position of the molecule at a later time t. In a macroscopically homogeneous system, on average, the probability of a displacement to the left is equal to a displacement to the right. Therefore the average displacement $< \mathbf{r}(t) - \mathbf{r}_0 >$ will be zero. Here, the brackets $< \cdots >$ denote thermal averaging, that is, averaging with respect to the probability that the position is equal to $\mathbf{r}$ at time t, given that the position is $\mathbf{r}_0$ at time $t = 0$ (we will give a precise definition of probability density functions later). The average displacement can therefore not be used to characterize the diffusive motion of a molecule. The most simple quantity that can be used for this purpose is the so-called *mean-squared displacement* $W(t)$, which is defined as,

$$W(t) \equiv <|\mathbf{r}(t) - \mathbf{r}_0|^2> \ . \tag{1}$$

Clearly this is a non-zero and non-trivial function of time.
Thermal motion of molecules leads to an overall net mass transport in case the concentration of the diffusing species varies with position (as depicted in Fig.1b for a solution of molecules, and in Fig.1c for relatively small molecules that diffuse through an essentially static environment of a crystalline solid). Mass is transported from the region of high concentration to the region of low concentration, as indicated by the arrows in Fig.1b,c. This type of diffusion process is referred to as **gradient diffusion** or **collective diffusion**, where "collective" refers to the coherent displacement of many molecules. An intuitive understanding of why net mass transport occurs due to concentration gradients is as follows. A given diffusing molecule experiences a different number of thermal collisions with neignbouring particles on the side facing the region with high concentration and that with low concentration. This results in a net force on that molecule, which therefore attains a net velocity. There are more collisions on the side where the concentration is high as compared to the side where the concentration is low, so that mass transport will typically occur towards regions of lower concentration.
A third type of diffusion process is the mixing of two (or more) molecular species, where the various species exhibit concentration gradients. Like for collective diffusion there is a net mass transport for each of the species, but the mass fluxes of the various species need not be in the

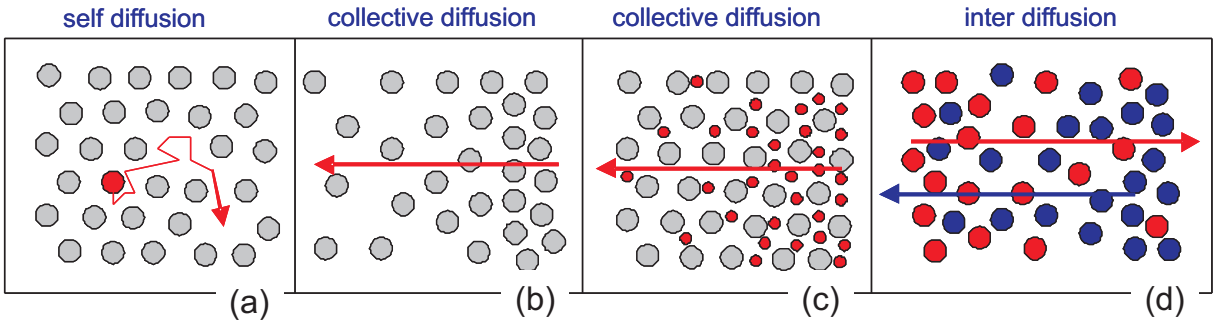


Fig. 1: *The three types of diffusion processes. (a) "self diffusion": the thermal motion of a single molecule, here for molecule within a crystalline solid with open interstitial positions. (b) "collective diffusion": a concentration gradient in a solution of molecules induced diffusive mass transport from the region of high concentration to the region of low concentration (as indicated by the arrow. (c) again collective diffusion, but now of small molecules (in red) through an essentially static environment of a crystalline solid. (d) "Inter diffusion": two different molecules (in blue and red), with opposite concentration gradients, mix due to diffusion. The arrows indicate the direction of net mass transport of the two components.*

same direction. This diffusion process is referred to as **inter diffusion**, and is depicted in Fig.1d, where the directions of the mass fluxes are indicated by the arrows.

Within the general classification of diffusion processes in self-, collective, and inter-diffusion, there is a great variety of different types of diffusion mechanisms for various types of systems. Some of these will be discussed quantitatively, and some only on a qualitative level. In this chapter we will focus on self- and collective-diffusion. Inter-diffusion will not be discussed.

Diffusive mass transport in stable systems is from regions of high concentration to low concentration. For thermodynamically unstable systems, however, strong attractive forces between molecules favor increase of concentration, where molecules are on average in each other's vicinity. The energy is lowered by increasing the concentration in part of the system due to the strong attractive interactions. Diffusion is now "uphill", from regions of low concentration to high concentration. Inhomogeneities thus increase in time, which is a kinetic stage during phase separation. The end-state is a coexistence between two phases. The initial stage of phase separation from an initially homogeneous, unstable system is discussed in section 4.

2 Self Diffusion

In this section we shall first develop a simple model in one dimension, where a molecule resides on discrete positions and can jump between these positions with a certain prescribed probability. In subsection 2.2 the discrete model will be cast in a continuum description, which allows for the explicit analysis of the time evolution of the probability density function for the position coordinate of a diffusing molecule. In subsection 2.3, the diffusion of a molecule in a periodic energy landscape will be analyzed, and compared to experiments. Finally, in subsection 2.4 a few other types of self-diffusion processes will be addressed on a qualitative level.

2.1 A simple model

As a first approach towards the understanding of diffusion processes, consider a single molecule that moves in one dimension and resides on discrete positions, as sketched in Fig.2. The distance between the discrete positions of the molecule is l, say, and the molecule is assumed to be at the site located at the origin at time $t=0$. The sites are indexed,

$$\{\cdots, -(n+1), -n, \cdots, -2, -1, 0, 1, 2, \cdots, n, (n+1), \cdots\} ,$$

where the "0" is the origin. The probability per unit time to make a "jump" to the left and right will be denoted by q and p, respectively. In the introduction we discussed self diffusion in case these two transition probabilities are equal. Taking $p \neq q$ is a generalization which can be thought of as self diffusion in an external force field, which induces a net average velocity of the molecule in the direction of the largest transition probability. We will come back to the effect of an external field at the end of this section.
Since diffusion is due to random thermal displacements, any theory that describes diffusion processes must be formulated in terms of probabilities. The probability to find the molecule at site n at time t will be denoted as $P(n,t)$. Since the probability to find the molecule at some site is unity, $P(n,t)$ is *normalized* in the sense that,

$$\sum_{n=-\infty}^{\infty} P(n,t) \;=\; 1 \; . \tag{2}$$

Since the molecule is supposed to be located at the origin at time zero, the probability for $n=0$ is unity at that time, while it is zero for all other n's,

$$P(n, t=0) \;=\; \delta_{n\,0} \; , \tag{3}$$

where the *Kronecker delta* $\delta_{n\,0}$ is unity for $n=0$ and zero for $n \neq 0$. For any function $f(n)$, its average value $< f > (t)$ at time t is equal to,

$$< f > (t) \;=\; \sum_{n=-\infty}^{\infty} f(n)\, P(n,t) \; . \tag{4}$$

An explicit expression for the probability $P(n,t)$ can in principle be found from the solution of its equation of motion, where the time derivative $\partial P(n,t)/\partial t$ is expressed in terms of $P(n,t)$. This so-called *master equation* can be constructed as follows. There is an increase of the probability to find the molecule at site n due to displacements from the neighbouring site $n-1$ to site n. The increase of $P(n,t)$ per unit time is equal to the probability $P(n-1,t)$ that the molecule is located at site $n-1$, multiplied by the transition rate p for the molecule to diffuse to the right. Similarly the increase of $P(n,t)$ per unit time due to jumps from site $n+1$ to the left is equal to $q\,P(n+1,t)$. There is a decrease of $P(n,t)$ due to jumps from site n to the neighbouring sites. This decrease per unit time is similarly equal to $(p+q)\,P(n,t)$. We thus arrive at the following master equation,

$$\frac{\partial}{\partial t} P(n,t) \;=\; p\,P(n-1,t) + q\,P(n+1,t) - (p+q)\,P(n,t) \; . \tag{5}$$

In principle this equation can be solved for $P(n,t)$, which then allows for the explicit calculation of averages (see eq.(4)).

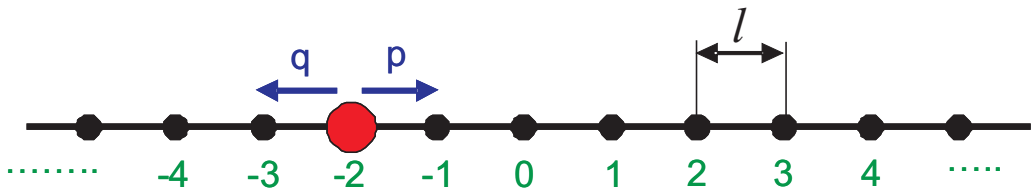


Fig. 2: *A molecule (indicated in red) diffusing in one dimension, which resides on lattice sites which are a distance l apart. At time zero the molecule is at the origin (the site with index 0). The transition rate for a displacement to the left is q, and to the right p.*

Of particular interest for self diffusion is the mean-squared displacement defined in eq.(1). In the extension to the case where $p \neq q$, as discussed above, there is an additional average that is related to the field-induced net displacement of the molecule, which we will refer to as the *drift velocity*. Let us first consider the average net velocity $<v>$ of the molecule, which is equal to,

$$<v> = l \frac{d}{dt} <n> (t) = l \sum_{n=-\infty}^{\infty} n \frac{\partial}{\partial t} P(n,t) . \tag{6}$$

Multiplying both sides of eq.(5) with n and summation over all n's leads to,

$$\frac{d}{dt} <n> (t) = \sum_{n=-\infty}^{\infty} n \left\{ p\, P(n-1,t) + q\, P(n+1,t) - (p+q)\, P(n,t) \right\} . \tag{7}$$

Each of the sums on the right hand-side can be explicitly evaluated. Consider as an example the first sum, which can be written as,

$$\sum_{n=-\infty}^{\infty} n\, P(n-1,t) = \sum_{m=-\infty}^{\infty} (m+1)\, P(m,t) = <n> (t) + 1 , \tag{8}$$

where the normalization identity (2) has been used. The remaining two sums can be evaluated similarly, leading to,

$$<v> = l \left\{ p - q \right\} . \tag{9}$$

Note that in case $p = q$, the drift velocity is zero, as it should. Next consider the mean-squared displacement. Multiplying both sides of the master equation (5) with n^2 and summing over all n's, it is found that.

$$\frac{d}{dt} W(t) = l^2 \sum_{n=-\infty}^{\infty} n^2 \left\{ p\, P(n-1,t) + q\, P(n+1,t) - (p+q)\, P(n,t) \right\} . \tag{10}$$

Similarly to the identity (8), the first sum can be written as,

$$\sum_{n=-\infty}^{\infty} n^2 P(n-1,t) = \sum_{m=-\infty}^{\infty} \left(m^2 + 2\,m + 1 \right) P(m,t) = <n^2> (t) + 2 <n> (t) + 1 , \tag{11}$$

and similarly for the two other contributions. This gives,

$$\frac{d}{dt} W(t) = l^2 \left(p + q \right) + 2\, l^2 \left(p - q \right)^2 t , \tag{12}$$

where it is used that $< n > (t) = l\,(p-q)\,t$, which follows from eq.(9) for the drift velocity. Integration with respect to time, and noting that $W(t=0)=0$ thus gives,

$$W(t) \;=\; l^2\;(\,p+q\,)\;t + l^2\;(\,p-q\,)^2\;t^2\;. \tag{13}$$

For pure diffusive motion where $p=q$, the mean-squared displacement therefore varies linearly with time. Typical distances $\sqrt{W(t)}$ over which a molecule diffuses during a time t thus vary like $\sqrt{t}$. Such a time dependence is typical for diffusive processes. The contribution $\sim t^2$ to the mean squared displacement in eq.(13) originates from the constant drift velocity. Note that the mean squared displacement $W_v(t)$ in a reference frame that moves along with the drift velocity is equal to,

$$W_v(t) \;\equiv\; l^2 < (\,n - < v > t\,)\;(\,n - < v > t\,) > = \; l^2\,(\,p+q\,)\;t\;, \tag{14}$$

which is the mean squared displacement for equal p and q.
A simple model for the difference between p and q due to an external field is most easily illustrated by considering a charged molecule in an external electric field. In an equilibrium system, the difference in the Boltzmann probability to find the molecule at two neighbouring sites is proportional to $\exp\{-\beta\,Q\,E\;l\}$, where $\beta = 1/k_BT$ (with k_B Boltzmann's constant, T the temperature), Q the charge carried by the molecule and E the electric field strength. This suggest the following form for the transition probabilities,

$$\begin{aligned} p &= \alpha\,\exp\left\{+\tfrac{1}{2}\,\beta\,Q\,E\,l\right\}\;, \\ q &= \alpha\,\exp\left\{-\tfrac{1}{2}\,\beta\,Q\,E\,l\right\}\;, \end{aligned} \tag{15}$$

where the electric field is chosen in positive direction, towards increasing index numbers. The prefactor α is the value of p and q in the absence of the field. For sufficiently small electric field strengths, where the Boltzmann exponents can be expanded to linear order in the electric field, it is thus found from eq.(9) that,

$$< v > = \;\frac{1}{\gamma}\;F\;, \tag{16}$$

where $F = Q\,E$ is the force exerted by the field on the charged molecule, and $\gamma = k_BT/\alpha\,l^2$ is a "*friction coefficient*". In the stationary state, where the drift velocity is constant, independent of time, the total average force on the molecule is zero. That is, the force on the molecule arising from "friction" with the surrounding matter must be equal to F in magnitude, but is opposite in sign. The friction force F_{fr} of the molecule with its surroundings is thus proportional to its velocity: $F_{fr} = -\,\gamma\,< v >$. The mean squared displacement, relative to the co-moving frame follows from eq.(14) as,

$$W_v(t) \;=\; 2\,D_s\,t\;, \tag{17}$$

where the **self diffusion coefficient** D_s is equal to,

$$D_s \;=\; \frac{k_B\,T}{\gamma}\;. \tag{18}$$

Since W_v is equal to the mean squared displacement in the absence of the field, this relation connects diffusive properties to the friction coefficient. This opens a way to calculate the self diffusion coefficient through the calculation of the friction coefficient. The relation (18) has been put forward by Einstein, and is therefore commonly referred to as the *Einstein relation*. This relation is generally valid, not just within the realm of the present simple model, as will be discussed later.

2.2 A continuum description

The master equation (5) can be cast into a differential equation, taking the limit where the distance l between the sites tends to zero. We consider here the case where the two transition probabilities p and q are both equal to α, say. To take the continuum limit, the master equation is rewritten as,

$$\frac{\partial}{\partial t} P(n,t) \;=\; \alpha\, l^2\, \frac{1}{l} \left[\frac{P(n+1,t) - P(n,t)}{l} - \frac{P(n,t) - P(n-1,t)}{l} \right] . \tag{19}$$

Each of the terms between the brackets is a first order spatial derivative with respect to position in the limit that $l \to 0$, so that the entire combination is a second order derivative. Replacing n by the continuously varying position x, it follows that,

$$\frac{\partial}{\partial t} P(x,t) \;=\; D_s \frac{d^2}{dx^2} P(x,t) \; , \tag{20}$$

where $D_s = \alpha\, l^2 = k_B T/\gamma$ is the self diffusion coefficient that was already introduced in the previous subsection. This equation of motion is easily extended to three dimensions, assuming that thermal displacements in the different Cartesian directions are independent, and the diffusion coefficient is the same for the three dimensions,

$$\frac{\partial}{\partial t} P(\mathbf{r},t) \;=\; D_s \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] P(\mathbf{r},t) \equiv D_s\, \nabla^2\, P(\mathbf{r},t) \; . \tag{21}$$

Here $\mathbf{r} = (x,y,z)$ is the position coordinate of the molecule in three-dimensional space. Equations of motion of this sort are commonly referred to as **diffusion equations**, of which more general forms will be discussed later.
Contrary to the discrete case, where $P(n,t)$ is the probability to find the diffusing molecule at site n, the probability to find the molecule at a given position $\mathbf{r}$ with infinite accuracy is zero. In the continuum limit we have to work with so-called *probability density functions* (pdf's). The above pdf $P(\mathbf{r},t)$ is to be understood as follows,

> $P(\mathbf{r},t)\, d\mathbf{r}$ is the probability to find the molecule at time t with its center of mass within the infinitesimally small volume element $d\mathbf{r} = dx\, dy\, dz$ that is located at $\mathbf{r}$.

The continuum analogue of a thermal average for a discrete variable in eq.(4) is now a sum over all infinitesimally small boxes, that is,

$$< f > (t) \;=\; \int d\mathbf{r}\; f(\mathbf{r})\; P(\mathbf{r},t) \; , \tag{22}$$

for any (well-behaved) function f. Similar to the previous subsection, the mean squared displacement can be calculated without having to solve the diffusion equation. Taking, without loss of generality, the molecule at the origin at time $t = 0$, the mean squared displacement is equal to (see eq.(1)),

$$W(t) \;=\; \int d\mathbf{r}\; r^2\, P(\mathbf{r},t) \; . \tag{23}$$

Multiplying both sides of the diffusion equation with r^2 and integration leads to,

$$\frac{d}{dt} W(t) \;=\; D_s \int d\mathbf{r}\; r^2\, \nabla^2 P(\mathbf{r},t) \;=\; D_s \int d\mathbf{r}\; P(\mathbf{r},t)\, \underbrace{\nabla^2 r^2}_{=6} \;=\; 6\, D_s \; . \tag{24}$$

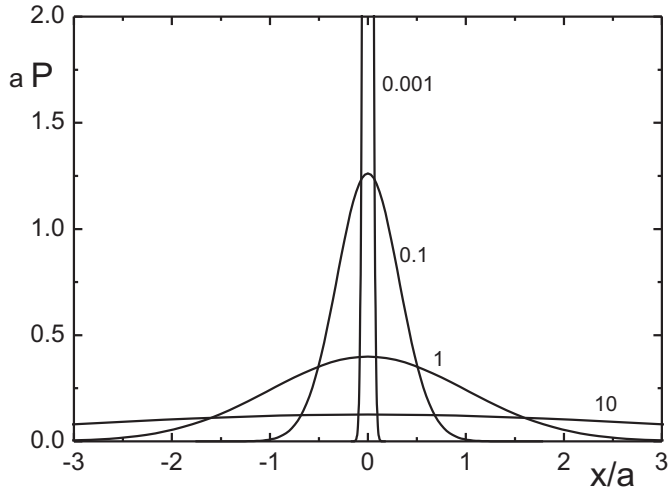


Fig. 3: *The probability density function $P(x,t)$ in eq.(27) (multiplied by the radius a of the spheres, as a function of x/a (with x the x-component of the position coordinate), for several values of the dimensionless time $\tau = t/\tau_a$, with $\tau_a = a^2/2\,D_s$ (the time required to reach a mean squared displacement in one direction equal to a^2), as indicated in the figure.*

In the second equation Green's second integral theorem is applied, and in the last line the continuum analogue of the normalization condition (2) is used. Note that each dimension contributes a $2D_s$. Since $W(t=0) = 0$ it follows that,

$$W(t) \;=\; 6\,D_s\,t\;, \tag{25}$$

in accordance with the result (17) for the mean square displacement in one dimension.
The initial condition that the molecule is at the origin at time $t = 0$ is mathematically formulated as,

$$P(\mathbf{r}, t=0) \;=\; \delta(\mathbf{r})\;, \tag{26}$$

where $\delta(\mathbf{r})$ is the *delta distribution*. This is the continuum analogue of the initial condition (3) for the discrete model. Imposing the above initial condition, the diffusion equation (21) can be solved analytically (with $r = |\,\mathbf{r}\,|$),

$$P(\mathbf{r},t) \;=\; \frac{1}{[\,4\,\pi\,D_s\,t\,]^{3/2}}\,\exp\left\{-\,\frac{r^2}{4\,D_s\,t}\right\}\;. \tag{27}$$

This function can be written as a product of the pdf for the x-, y- and z-coordinate, where the pdf for the x-coordinate is equal to,

$$P(x,t) \;=\; \frac{1}{[\,4\,\pi\,D_s\,t\,]^{1/2}}\,\exp\left\{-\,\frac{x^2}{4\,D_s\,t}\right\}\;, \tag{28}$$

and similarly for the y- and z-coordinates. This is a Gaussian pdf with a width that increases with time, as depicted in Fig.3. For short times the pdf is very sharply peaked (and ultimately

approaches the delta distribution as $t \rightarrow 0$), while for larger times the probability for large distances increases due to the diffusive displacement of the tracer molecule.
The equation of motion (21) can be generalized to include an external force acting on the molecule. First notice that the probability $P(\mathbf{r},t)$ can in principle be measured as follows. A molecule that resides at the origin is released at time zero, after which its trajectory is recorded. This experiment is repeated many times. In each of the experiments, the molecule follows a different trajectory. The number of trajectories that intersect, at time t, a small box centered at $\mathbf{r}$ measures the pdf $P(\mathbf{r},t)$. Alternatively, many non-interacting molecules can be released simultaneously, and each of the trajectories is recorded. Since the molecules do not interact with each other, they diffuse after release at time zero as if they were alone in the system. Each of the molecules thus exhibit self diffusion as described above. Again, the number of trajectories that intersect a small box at $\mathbf{r}$ after a time t measures the pdf $P(\mathbf{r},t)$. That number of trajectories, however, is also proportional to the local concentration $\rho_0(\mathbf{r},t)$ that one would measure (where the index "0" refers to non-interacting molecules). We can thus interpret $P(\mathbf{r},t)$ as the concentration $\rho_0(\mathbf{r},t)$ that exists when many non-interacting molecules are released at the origin $\mathbf{r}=\mathbf{0}$ at time $t=0$. The number density obeys the exact continuity equation,

$$\frac{\partial}{\partial t}\,\rho_0(\mathbf{r},t) \;=\; -\,\nabla\cdot\left[\,\rho_0(\mathbf{r},t)\,\mathbf{v}(\mathbf{r},t)\,\right]\;, \tag{29}$$

where $\mathbf{v}(\mathbf{r},t)$ is the thermally averaged velocity of molecules. When inertial forces are neglected (we shall comment on this at the end of this section), according to Newton's equation of motion, there is a balance of forces, that is, all non-inertial forces add up to zero. There are three forces to be considered. First of all there is the friction force $\mathbf{F}^{fr}=-\gamma\,\mathbf{v}$ that arises from interactions of the molecule with surrounding matter, where γ is the friction coefficient that was already introduced at the end of subsection 2.1. The second force is responsible for the velocity that the molecule attains in the absence of an external force field, and is referred to as the **Brownian force** $\mathbf{F}^{Br}$. The third force $\mathbf{F}^{ext}$ is due to an external field. By force balance we have,

$$\mathbf{F}^{fr}+\mathbf{F}^{Br}+\mathbf{F}^{ext} \;=\; \mathbf{0} \qquad\rightarrow\qquad \mathbf{v} \;=\; \frac{1}{\gamma}\,\left[\,\mathbf{F}^{Br}+\mathbf{F}^{ext}\,\right]\;. \tag{30}$$

Substitution into eq.(29) and comparing with eq.(21), with P replaced by ρ_0, in the absence of an external field, leads to $\mathbf{F}^{Br}=-\gamma D_s\,\nabla\ln\{\,\rho_0(\mathbf{r},t)\,\}$. For a conservative force field, for which $\mathbf{F}^{ext}=-\nabla\Phi$ (with Φ the external potential), and in case equilibrium is reached, the density must be proportional to the Boltzmann exponential $\rho_0\sim\exp\{-\Phi/k_BT\}$. On the other hand $\mathbf{v}=\mathbf{0}$ in equilibrium, which immediately leads to $\gamma\,D_s\,\nabla\ln\{\rho_0\}+\nabla\Phi=\mathbf{0}$, and hence, $\rho_0\sim\exp\{-\Phi/\gamma\,D_s\}$. It follows that $D_s=k_BT/\gamma$, which reproduces the Einstein relation (18) that we found within the simple model considered in subsection 2.1. It also follows that the Brownian force is equal to,

$$\mathbf{F}^{Br}(\mathbf{r},t) \;=\; -k_B\,T\,\nabla\,\ln\{\rho_0(\mathbf{r},t)\}\;, \tag{31}$$

with ρ_0 replaced by $P(\mathbf{r},t)$ when used in the equation of motion for the pdf $P(\mathbf{r},t)$. Using this expression for the Brownian force in eq.(29), and replacing the density ρ_0 by the pdf $P(\mathbf{r},t)$ thus leads to the generalization of the equation of motion eq.(21) to include an external force field,

$$\frac{\partial}{\partial t}\,P(\mathbf{r},t) \;=\; D_s\,\nabla\cdot\left[\,\nabla P(\mathbf{r},t)-\beta\,P(\mathbf{r},t)\,\mathbf{F}^{ext}(\mathbf{r})\,\right]\;. \tag{32}$$

We will use this equation in the next subsection to calculate the self diffusion coefficient of a molecule in a prescribed energy landscape.
In the above we assumed that inertial forces can be neglected. This is allowed on a time scale where the thermally averaged momentum coordinate, with a given initial value, relaxes to zero. According to Newton's equation of motion, without an external field,

$$\frac{d\,\mathbf{p}(t)}{dt} \;=\; -\frac{\gamma}{m}\,\mathbf{p}(t)\;, \tag{33}$$

where $\mathbf{p} = m\,\mathbf{v}$ is the momentum coordinate and m is the mass of the molecule. It follows that (with $\mathbf{p}_0$ the initial momentum at time zero),

$$\mathbf{p}(t) \;=\; \mathbf{p}_0\;\exp\{-t/\tau\} \quad , \quad \tau \;=\; m/\gamma\;. \tag{34}$$

The time scale τ on which the momentum coordinate relaxes is very typically much smaller than the time during which appreciable diffusion occurs, which validates the neglect of inertia. The time scale on which inertial forces can be neglected is commonly referred to as the **diffusive time scale**, while the dynamics on this time scale is referred to as **overdamped dynamics**. "overdamped" refers to the fact that friction forces are much larger than inertial forces.

2.3 Diffusion through a one-dimensional periodic energy landscape

As an example of an explicit calculation of the self diffusion coefficient we consider a molecule that interacts with its surroundings as described by a potential energy Φ. This potential is assumed to be periodic in one dimension, say the z-direction, and is constant along the other two directions : $\Phi \equiv \Phi(z)$. An experimental example that will be discussed at the end of this section is a rod-like molecule in a smectic phase that diffuses from one smectic layer to the other. The potential is assumed to be periodic with a period l, that is, $\Phi(z) = \Phi(z + n\,l)$ for any integer n. We shall calculate the friction coefficient γ and employ the Einstein relation $D_s = k_B T/\gamma$ to obtain the self diffusion coefficient in terms of the potential. In addition to the force $-\nabla\Phi$ due to interactions with the surroundings, there is thus an additional applied constant force F^{app} in the z-direction on the molecule that leads to a finite thermally averaged velocity.
For short times, the molecule "rattles within potential valleys". For longer times the molecule moves from one valley to the other, that is, it moves across potential barriers. On can therefore distinguish between a **short-time self-diffusion coefficient** that describes the thermal motion within a potential valley in the z-direction, and the **long-time self diffusion coefficient** that describes thermal displacements between valleys. Here we calculate the long-time self diffusion coefficient, which is relevant to mass transport on larger length scales.
The long-time self diffusion coefficient can be calculated from an eigen-function expansion of the solution of the diffusion equation (32), with $\mathbf{F}^{ext} \equiv -\nabla\Phi + \mathbf{F}^{app}$ [1]. Alternatively, an expression for the long-time self diffusion coefficient can be obtained without having to solve the diffusion equation explicitly [2]. First of all, we identify the pdf $P(x,t)$ as the concentration of non-interacting molecules, an identification that has been discussed in subsection 2.2. Instead of analyzing the motion of a single molecule under the action of the force $\mathbf{F}^{app}$, one can analyze the motion of many non-interacting molecules simultaneously. Since the molecules do not interact with each other, they move as if they were alone in the system, and hence they all move like a self-diffuser. The stationary flux j_0 of molecules in the x-direction is equal to (the indices

"0" refer to "non-interacting molecules"),

$$j_0 \;=\; \rho_0(z)\, v(z) \;=\; \bar{\rho}\, \exp\{-\beta\, \Phi(z)\}\; v(z)\;, \tag{35}$$

where $\bar{\rho}$ is the number density that would have existed in the absence of the potential. This expression is valid up to order $(F^{app})^2$. The velocity is linear in F^{app}, so that the zeroth-order solution for $\rho_0 \equiv P$ of eq.(32) with $\mathbf{F}^{ext} \equiv -\nabla\Phi$ can be used.
Let τ be the mean time spend by a molecule within a valley. The mean velocity $\bar{v}$ can then be expressed as (again, l is the periodicity of the potential),

$$\bar{v} \;=\; l/\tau\;. \tag{36}$$

The number of molecules N_u within a single valley over an area $A_\perp$ in the yz-plane is equal to,

$$N_u/A_\perp \;=\; j_0\,\tau \;=\; \int_{-l/2}^{l/2} dz\; \rho_0(z) \;=\; l\,\bar{\rho}\;\prec \exp\{-\beta\,\Phi\}\succ\;, \tag{37}$$

where the brackets define the periodicity average,

$$\prec f \succ \equiv\; \frac{1}{l}\int_{-l/2}^{l/2} dz\; f(z)\;, \tag{38}$$

for any function $f(z)$. The last step in eq.(37) is valid up to leading order in F^{app}. Note that j_0 is linear in F^{app}, while τ varies like $1/F^{app}$ for sufficiently small applied forces, so that their product is a constant to leading order.
The force balance relation (30) for the present case reads,

$$F^{fr} + F^{app} - \frac{d}{dz}\,\{\, k_B T \ln\rho_0 + \Phi\,\} \;=\; 0\;. \tag{39}$$

The local friction force F^{fr} is equal to $-\gamma_0\, v(z)$, where γ_0 is the friction coefficient in the absence of the potential Φ. On integration of both sides from $z = -l/2$ to $+l/2$, the gradient contribution vanishes due to symmetry, and hence,

$$F^{app} \;=\; \gamma_0\,\frac{1}{l}\int_{-l/2}^{l/2} dz\; v(z)\;. \tag{40}$$

Combining the above equations we have,

$$\begin{aligned}
\frac{1}{l}\int_{-l/2}^{l/2} dz\; v(z) \;&\underbrace{=}_{eq.(35)}\; \frac{j_0}{\bar{\rho}}\;\prec \exp\{+\beta\Phi\}\succ \\
&\underbrace{=}_{eq.(37)}\; \frac{l}{\tau}\;\prec \exp\{-\beta\Phi\}\succ \prec \exp\{+\beta\Phi\}\succ \\
&\underbrace{=}_{eq.(36)}\; \bar{v}\;\prec \exp\{-\beta\Phi\}\succ \prec \exp\{+\beta\Phi\}\succ\;.
\end{aligned} \tag{41}$$

Hence, by substitution into eq.(40),

$$F^{app} \;=\; \gamma\,\bar{v}\quad,\quad \text{with}\quad \gamma \;=\; \gamma_0\;\prec \exp\{-\beta\Phi\}\succ \prec \exp\{+\beta\Phi\}\succ\;. \tag{42}$$

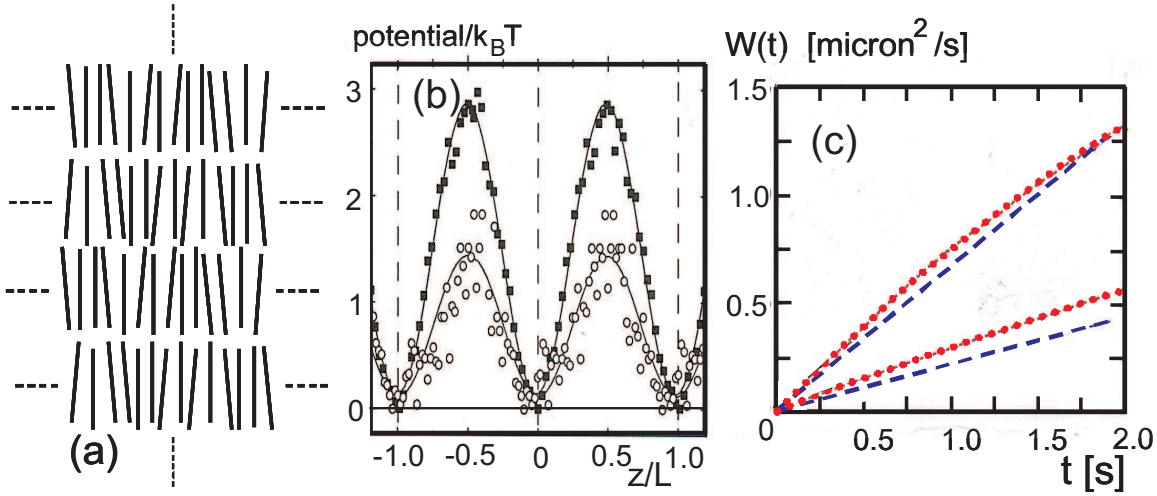


Fig. 4: *(a) A sketch of a smectic phase. (b) The potential that a given rod experiences due to interactions with surrounding rods. Two potentials are shown: the solid data points* ■ *are for an ionic strength of* $20\ mM$*, and the open symbols* ∘ *refer to an ionic strength of* $110\ mM$*. (c) Mean-squared displacements as a function of time for the two ionic strengths* 20 *(lower two curves) and* $110\ mM$ *(upper curves). The red dotted lines are experimental, and the blue dashed lines are obtained from eq.(43). Data are taken from Ref.[3].*

The long-time self diffusion coefficient D_s^l is thus found from the Einstein relation to be equal to,

$$D_s^l \;=\; \frac{D_s^0}{\prec \exp\{-\beta\Phi\} \succ \prec \exp\{+\beta\Phi\} \succ}\;, \tag{43}$$

where D_s^0 is the self diffusion coefficient in the absence of the potential Φ.
Diffusion of colloidal rod-like particles across smectic layers provides an experimental test of this relation [3, 4]. A smectic phase is spontaneously formed by systems of rod-like particles at sufficiently high concentration. This structure consists of a stack of mono-layers of rods with a preferred direction of alignment in the stacking direction, as sketched in Fig.4a. There is diffusion within the layers, and there is an exchange of rod-like particles between the layers. The long-time self diffusion coefficient corresponding to exchange between layers is given by eq.(43), where the potential is now set up by the interaction of the tracer rod with the surrounding rods. The assumption here is that the fluctuations of the position and width of the smectic layers can be neglected, so that the potential is a given function of the position of the tracer rod. This potential can be measured through the residence time of rods at prescribed positions. The residence time is inversely proportional to the probability $\sim \exp\{-\beta\,\Phi(z)\}$ to find a rod at position z (the direction perpendicular to the smectic planes). The experiments discussed here are performed with fd-virus particles in water. Fd-viruses are very long ($880\ nm$) and thin ($6.8\ nm$) stiff rods, consisting of a DNA strand which is rendered stiff by so-called coat-proteins that are attached to the DNA. The potential for this system, as obtained from residence times, is given in Fig.4b for two ionic strengths: 20 and $110\ mM$. The ionic strength changes the interactions between the rods and therefore the potential set up by the smectic layers. The solid line is a best fit to a sinusoidal function, which described the potential accurately. The long-time self diffusion coefficient can be measured independently by tracing rod-like colloids by means of time-resolved microscopy, where the tracer rod is fluorescently labeled while the

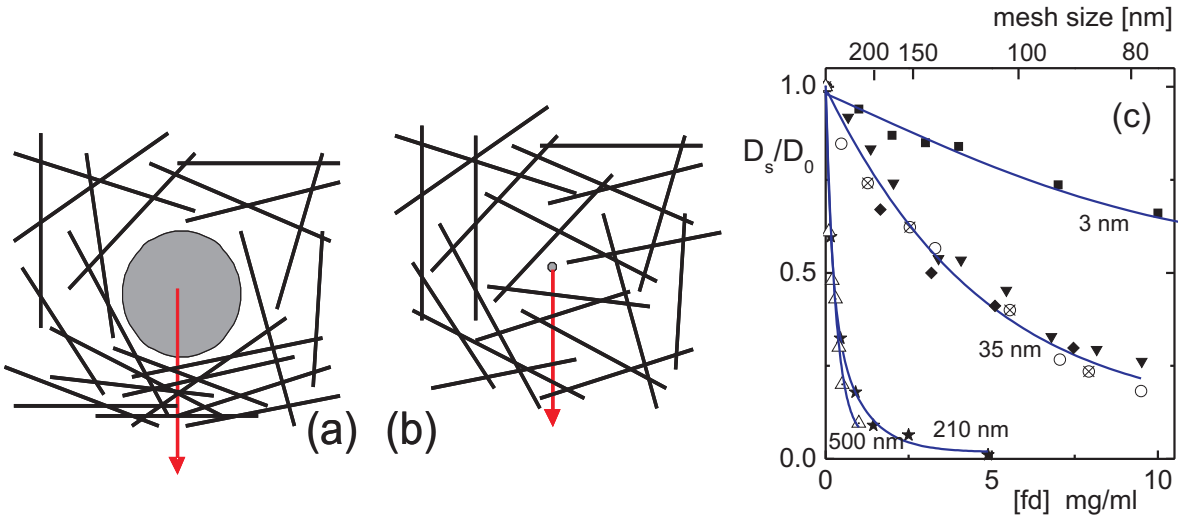


Fig. 5: *(a) Diffusion of a sphere through a rod-network, where the sphere is much large than the mesh size of the network. A displacement is accompanied by structural deformation of the network. Typically, rods accumulate in front of the sphere. (b) Diffusion of a small sphere, much smaller than the mesh size. The network structure is essentially unaffected by motion of the sphere, and hydrodynamic interactions with the network are dominant. (c) The long-time collective diffusion coefficient of spherical tracers of various sizes (as indicated) through fd-virus fiber networks as a function of fd-concentration (lower axis) and the mesh size of the network (upper axis). Various symbols of data points refer to different experimental techniques (particle tracking, fluorescence correlation spectroscopy and dynamic light scattering). This plot is taken from Ref.[6].*

remaining rods are not fluorescent. The resulting mean-squared displacement is given as a function of time in Fig.4c, for the two ionic strengths (the red dotted lines). The blue dashed lines in Fig.4c corresponds to the prediction (43) with the use of the potential given in Fig.4a. The slope of these lines line is twice the diffusion coefficient, since we are dealing here with diffusion in one dimension. The free diffusion coefficient D_s^0 is in this case the diffusion coefficient along the director in the nematic phase, just below the nematic-smectic phase transition concentration. The comparison of the experiments with the prediction in eq.(43) involves therefore no fitting parameters. The experimental upper curves deviate somewhat from straight lines, which is probably due to the transition from short-time diffusion to long-time diffusion. Quite detailed experiments on diffusion of spherical colloids through a sinusoidally varying energy landscape set up by an electric field can be found in Ref.[5].

2.4 Other self-diffusion processes

There are many other types of self-diffusion processes. A concise overview of all the self-diffusion processes and mechanisms is outside the scope of this chapter. In this subsection, we will discuss three other self-diffusion processes on a qualitative level.
Diffusion of spherical colloids through networks of very long and thin rods (or fibers) is another example where the friction coefficient is affected by interactions of the diffusion species (the

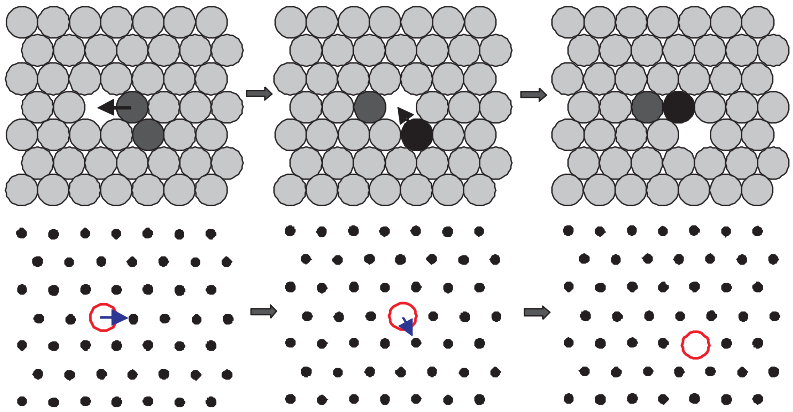


Fig. 6: *The mechanism of vacancy diffusion. In the upper three panels, the consecutive nearest neighbour displacement of the black-coloured molecules lead to motion of the vacancy, as shown schematically by the red circle in the lower panel.*

tracer sphere) and its surroundings (the rod-network). Such a self-diffusion process can not be described in terms of an external potential like for the above example in subsection 2.3. For this case the diffusion equation must be solved explicitly, which, again, is beyond the scope of this chapter. In case the tracer sphere sphere is much larger than the mesh size of the rod-network, the friction coefficient is mainly affected by the deformation of the network as the sphere is pulled through (see the sketch in Fig.5a). At first sight one might expect that diffusion of very small spheres, much smaller than the mesh size of the network (as sketched in Fig.5b) is essentially equal to the free diffusion diffusion coefficient, in the absence of the network. There is, however, an aspect that we have not discussed so far, which affects the friction coefficient even for these small tracer spheres. As the sphere moves through the solvent, it sets the solvent in motion. This fluid flow will be reflected by the network back to the sphere, which is thereby affected in its motion. These so-called **hydrodynamic interactions** contribute to the (long-time) friction coefficient, so that the diffusion coefficient is less than that of the freely diffusing sphere. For the large sphere, hydrodynamic interactions are relatively small as compared to direct interactions with the rods, while for small spheres, with a diameter that is of the order or smaller than the network mesh size, hydrodynamic interactions are dominant. Experimental data for long-time diffusion are given in Fig.5c, where the fiber network is formed by the same fd-virus particles as discussed in subsection 2.3 (which are $880\ nm$ long and $6.8\ nm$ thick, and relatively stiff). As can be seen, the effect of deformation of the network for the large spheres has a much more pronounced effect that hydrodynamic interactions for the smaller spheres. More details can be found in Refs.[6, 7, 8].

In crystals the position of single vacancies changes due to thermal motion of the surrounding molecules. The thermal displacement of a vacancy is the result of motion of neighbouring molecules to the actual vacancy position (as schematically shown in Fig.6). As the vacancy displacement is related to thermal motion of neighbouring molecules, vacancies obey the same diffusion laws as if it were a material-particle, and one can correspondingly define a **vacancy diffusion coefficient**. In order for the vacancy to move to a neighbouring position, the corresponding motion of a molecule that moves in the opposite direction requires it to move over an energy barrier (of height E, say) that is determined by the interactions with the remaining molecules in the crystal. Since the probability for a molecule to attain this energy is proportional to $\exp\{-\beta E\}$, the vacancy diffusion coefficient is also proportional to this exponent.

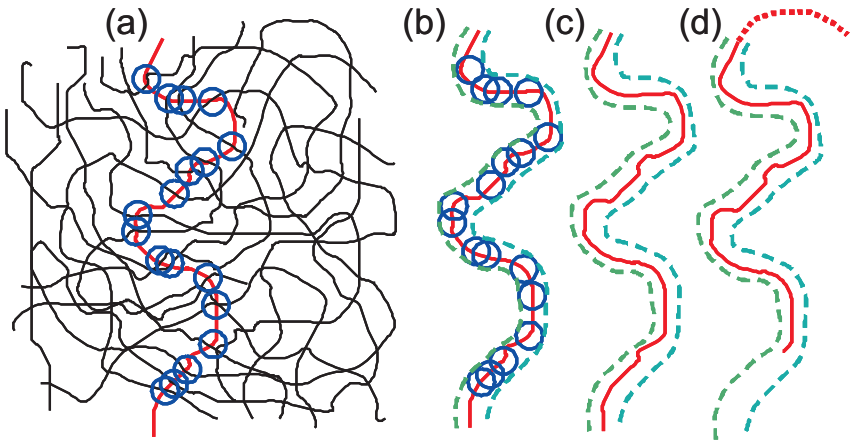


Fig. 7: *Reptation diffusion mechanism of flexible polymers. (a) a polymer chain (in red) within the matrix of other chains. The interaction points with other chains (indicated by circles in (b)) creates a "tube" as indicated in green in (c), through which the tagged chain can move out by diffusion, as depicted in (d).*

The same arguments hold for the diffusion of an interstitial atom through a crystal. Experimentally one therefore often finds that self diffusion coefficients exhibit a so-called **Arrhenius behaviour**, that is, the logarithm of the diffusion coefficient varies linearly with the reciprocal temperature $1/T$. The slope is equal to $-E/k_B$, from which the "**diffusion activation energy**" E is obtained.
So far only molecules of a spherical- and rod-like molecules have been discussed. Flexible molecules, like polymers, will of course also exhibit diffusion. Due to the high degree of entanglement in melts and solutions of polymers, single-polymer diffusive motion can be described by the so-called **tube model**. The interactions with neighbouring polymers (see Fig.7a) defines a "tube" within which a given polymer chain can diffuse (see Figs.7b,c). The given polymer chain can then find another tube after diffusion of one of the ends of the polymer chain outward the original tube (see Fig.7d) [9, 10, 11]. This **reptation mechanism** has been extended to include, for example, the dynamics of the tube itself and the retraction of the polymer within the tube.

3 Collective Diffusion

Contrary to self diffusion, collective diffusion describes the net mass transport due to gradients in concentration, as already discussed in the introduction. Similar to self diffusion, the first step towards a theory to describe diffusive mass transport is to derive an equation of motion for the concentration of the diffusing species. A simple diffusion equation has been proposed by Fick more than a century ago. The assumption he made is that the mass flux $\mathbf{j}$ of molecules is linear in concentrations gradients, that is,

$$\mathbf{j}(\mathbf{r},t) \;=\; -D_c\,\nabla\rho(\mathbf{r},t)\;, \tag{44}$$

provided that concentration gradients are sufficiently small. Here $\rho(\mathbf{r},t)$ is the instantaneous number density (number of molecules per unit volume) at position $\mathbf{r}$ and time t. Furthermore, the proportionality constant D_c is referred to as the **collective diffusion coefficient**, and $\nabla =$

$(\partial/\partial x, \partial/\partial y, \partial/\partial z)$ is the nabla- or gradient-operator. A minus sign is added to the right hand-side in eq.(44) to render D_c positive (note that the flux is typically directed towards regions of low concentration, as already discussed in the introduction). Substitution of this expression for the flux into the continuity equation gives,

$$\frac{\partial}{\partial t}\,\rho(\mathbf{r},t) \;=\; -\nabla\cdot\mathbf{j}(\mathbf{r},t) \;=\; \nabla\cdot\left[\,D_c\,\nabla\,\rho(\mathbf{r},t)\,\right]\;. \tag{45}$$

In concentrated systems, the collective diffusion coefficient depends on concentration, $D_c \equiv D_c(\rho(\mathbf{r},t))$, so that it can not placed in front of the $\nabla-$ operator. Suppose, however, that the deviation of the density from its average value is small. That is, $\rho(\mathbf{r},t) = \bar{\rho} + \delta\rho(\mathbf{r},t)$, with $\delta\rho(\mathbf{r},t)/\bar{\rho} \ll 1$, where $\bar{\rho}$ is the number density of the system without concentration gradients. When this is assumed, the diffusion equation (45) can be written, up to linear order in $\delta\rho(\mathbf{r},t)$, as,

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t} \;=\; D_c\,\nabla^2\rho(\mathbf{r},t)\;, \tag{46}$$

The assumption of small overall deviations from the average density is a quite severe assumption that is not satisfied for many cases, so that eq.(45) is relevant rather than eq.(46). For very dilute systems, where inter-molecular interactions are absent, this is a valid procedure, since then D_c is indeed a constant. In that case eq(46) reproduces eq.(21) for self diffusion (as discussed in subsection 2.2, this equation also holds for the concentration when interactions between the tracer molecules are absent). It follows that, for very dilute system, the self- and collective diffusion coefficients are identical,

$$D_s^0 \;=\; D_c^0\;, \tag{47}$$

where the superscript "0" is used to indicate that interactions between diffusing species are absent. In most practical systems, the concentration of diffusing molecules is large, such that inter-molecular interactions are important for the diffusive properties. The collective diffusion coefficient under such conditions is different from D_c^0, and depends on the interactions between molecules. We will therefore extend the diffusion equation for infinite dilution to include inter-molecular interactions in subsection 3.1. In subsection 3.2, **Fick's law** (45) will be derived from this generalized diffusion equation, and an explicit expression will be obtained for the collective diffusion coefficient in terms of the interaction potential between the molecules. The specific example of diffusion of hydrogen through metal crystals is then discussed in subsection 3.3.
One possible way to think about the physical meaning of the collective diffusion is as follows. Suppose that at time $t = 0$ there is a sinusoidal concentration profile, $\rho(\mathbf{r}, t=0) = \bar{\rho} + \Delta\rho_0 \times \sin\{\mathbf{k}\cdot\mathbf{r}\}$, where $\Delta\rho_0$ is the initial amplitude of the density variation superimposed on a constant overall concentration $\bar{\rho}$. The wavelength Λ of this sinusoidal density variation is equal to,

$$\Lambda \;=\; \frac{2\,\pi}{k}\;, \tag{48}$$

with k the length of the *wave vector* $\mathbf{k}$. Substitution of the Ansatz $\rho(\mathbf{r},t) = \bar{\rho} + \Delta\rho(t)\times\sin\{\mathbf{k}\cdot\mathbf{r}\}$ into Fick's diffusion equation (46) gives the time-dependent amplitude of the density variation equal to,

$$\Delta\rho(t) \;=\; \Delta\rho_0\,\exp\{-D_c\,k^2\,t\}\;. \tag{49}$$

The collective diffusion coefficient thus determines how fast a sinusoidal concentration profile relaxes. The relaxation rate varies with the wavelength as $\sim \Lambda^{-2}$, the interpretation of which is that it takes particles longer to diffuse over longer distances, while the time it takes to diffuse over a certain distance is quadratically depending on that distance (as quantified by eq.(25) for the mean-squared displacement of a single particle). As will be seen in the next subsection, the validity of eq.(46) (and hence of eq.(49)) is limited to wavelengths that are much larger than the range of the pair-interaction potential between the diffusing particles. When the pair-interactions between particles is repulsive, it is intuitively obvious that the particles in regions of high concentration are pushed apart more strongly at higher overall concentrations $\bar{\rho}$, which leads to a faster relaxation. In other words, the collective diffusion coefficient is expected to increase with increasing overall concentration. For the same reason a decrease is expected for particles with attractive pair-interaction potentials. The above intuitive arguments are only valid for sufficiently low concentrations $\bar{\rho}$, such that interactions between two given particles is not too much affected by the presence of other particles. At sufficiently high concentrations, the collective diffusion coefficient can decrease with increasing concentration $\bar{\rho}$ due to indirect interactions at large overall concentrations also for repulsive pair-interactions potentials. In addition to direct interactions, also hydrodynamic interactions can play an important role in the concentration dependence of the collective diffusion coefficient of particles in a solvent. A moving particle in a solvent induces a fluid flow that affects other particles in their motion, and therefore the value of the collective diffusion coefficient. Such hydrodynamic interactions will not be discussed in this chapter (more about the concentration dependence of diffusion coefficients of spherical particles in a solvent and hydrodynamic interactions can be found in, for example, Ref.[12, 13, 14]). While the collective diffusion coefficient increases with concentration in case of repulsive interactions, it is obvious that the self diffusion coefficient decreases. Due to the repulsive interactions with neighbouring particles, the diffusive displacement of a given particle is hindered, leading to a decrease of the mean-squared displacement, and thus to a smaller self diffusion coefficient.
Attractive interactions between particles can be so strong, that it is energetically favorable to increase an initial inhomogeneity. According to eq.(49) this happens when the collective diffusion coefficient is negative. A homogeneous system with a negative collective diffusion coefficient is unstable in the sense that arbitrary small fluctuations in the concentration leads to the formation of inhomogeneities. Such a growth of inhomogeneities, that ultimately leads complete phase separation, is referred to as **spinodal decomposition**. Spinodal decomposition will be discussed in some detail in section 4.

3.1 A generalized diffusion equation

As discussed at the end of subsection 2.2, there is force balance in the overdamped limit (or equivalently, on the diffusive time scale). Since inertial forces can be neglected, all other forces add up to zero on the diffusive time scale. In the dilute limit, where interactions between molecules can be neglected, force balance results in the expression (30) for the velocity of a given molecule, with the Brownian force being given in eq.(31) (where the number density is now denoted as ρ, without the subscript "0", since we here consider the more general case of concentrated systems). For interacting molecules, at relatively high concentration, the external force in eq.(30) is equal to the force that acts on a given molecule due to interactions with neighbouring molecules. We will denote this force by $\mathbf{F}^I$ instead of $\mathbf{F}^{ext}$, to indicate that the force is now due to inter-molecular interactions.

The force $\mathbf{F}^I$ on a molecule at $\mathbf{r}$ due to the presence of a second molecule at $\mathbf{r}'$ is equal to $-\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|)$, with V the pair-interaction potential. The force experienced by the molecule at $\mathbf{r}$ is equal to this pair-force, multiplied by the number of neighbouring molecules around the give molecule at position $\mathbf{r}$, averaged over all positions $\mathbf{r}'$ of neighbouring molecules,

$$\mathbf{F}^I(\mathbf{r},t) \;=\; -\int d\mathbf{r}'\; g(\mathbf{r},\mathbf{r}',t)\,\rho(\mathbf{r}',t)\,\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|)\;, \tag{50}$$

where $g(\mathbf{r},\mathbf{r}',t)$ is the probability to find a molecule at $\mathbf{r}'$, given that there is a molecule at $\mathbf{r}$. This **conditional probability density function** is commonly referred to as **the pair-correlation function**. From the above mentioned force balance, we have, $\gamma_0\,\mathbf{v} = \mathbf{F}^{Br} + \mathbf{F}^I$, where γ_0 is the friction coefficient at infinite dilution, that is, in the absence of inter-molecular interactions. The corresponding flux $\mathbf{j} = \rho\,\mathbf{v}$ is thus equal to,

$$\mathbf{j}(\mathbf{r},t) \;=\; -D_c^0\left[\nabla\rho(\mathbf{r},t) + \beta\,\rho(\mathbf{r},t)\int d\mathbf{r}'\; g(\mathbf{r},\mathbf{r}',t)\,\rho(\mathbf{r}',t)\,\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|)\right]\;, \tag{51}$$

where, as before, $\beta = 1/k_BT$ and $D_c^0 = k_BT/\gamma_0$. This expression for the flux can be substituted into the exact conservation equation, to obtain a **generalized diffusion equation**,

$$\begin{aligned}\frac{\partial\rho(\mathbf{r},t)}{\partial t} \;&=\; -\nabla\cdot\mathbf{j}(\mathbf{r},t)\\ &=\; D_c^0\left[\nabla^2\rho(\mathbf{r},t) + \beta\,\nabla\cdot\left\{\rho(\mathbf{r},t)\int d\mathbf{r}'\; g(\mathbf{r},\mathbf{r}',t)\,\rho(\mathbf{r}',t)\,\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|)\right\}\right]\;.\end{aligned} \tag{52}$$

This is the fundamental equation of motion that will be used in the sequel to analyze the collective diffusion at finite concentrations.

3.2 Derivation of Fick's law for concentrated systems

The generalized diffusion equation (52) can be used to derive Fick's law (45), where an explicit expression will be obtained for the collective diffusion coefficient D_c in terms of the pair-interaction potential V. The linear relationship (44) between the mass flux and spatial gradients in the concentration is expected to hold only for sufficiently small gradients. We will therefore assume that on the distance R_V over which the pair-potential V falls off to zero, the density is essentially a linear function of position. The following Taylor expansion can thus be used in the integral in eq.(52) (with $\mathbf{R} = \mathbf{r}^{'} - \mathbf{r}$,),

$$\rho(\mathbf{r}^{'},t) \;\approx\; \rho(\mathbf{r},t) + \mathbf{R}\cdot\nabla\rho(\mathbf{r},t)\;, \tag{53}$$

since the pair-potential limits the integration range in $\mathbf{r}^{'}$-space to a region of extent R_V around $\mathbf{r}$. Here the "$\cdot$" is the inner product of the two vectors on both sides. An appropriate approximation for the pair-correlation function can be obtained as follows. First of all, we need an approximation only for $|\,\mathbf{r}-\mathbf{r}^{'}\,| \leq R_V$. For such small, microscopic distances, the pair-correlation function relaxes essentially instantaneous to equilibrium relative to the time scale on which the concentration evolves. We can therefore take the pair-correlation function equal to its equilibrium value g^{eq}. This approximation can be considered as a statistical mechanical analogue of the **thermodynamic local equilibrium** assumption that is made in the theory of

irreversible thermodynamics. A natural choice is to take this equilibrium function equal to that of a homogenous system with a density in between the points $\mathbf{r}$ and $\mathbf{r}^{'}$,

$$g(\mathbf{r},\mathbf{r}^{'},t) \;\approx\; g^{eq}(|\,\mathbf{r}-\mathbf{r}^{'}\,|) \quad , \quad \rho \equiv \rho\left(\frac{\mathbf{r}+\mathbf{r}^{'}}{2},t\right) \; . \tag{54}$$

Note that due to translational and rotational invariance of a homogeneous system, the pair-correlation function depends on $\mathbf{r}$ and $\mathbf{r}^{'}$ only through $|\,\mathbf{r}-\mathbf{r}^{'}\,|$. Thus, again Tayloring up to linear spatial dependencies, after writing $(\mathbf{r}+\mathbf{r}^{'})/2$ as $\mathbf{r}+\frac{1}{2}\,\mathbf{R}$,

$$\begin{aligned} g(\mathbf{r},\mathbf{r}^{'},t) \;&\approx\; g^{eq}(R) + \left\{\,\rho\left(\frac{\mathbf{r}+\mathbf{r}^{'}}{2},t\right) - \rho(\mathbf{r},t)\,\right\}\,\frac{d}{d\rho}\,g^{eq}(R) \\ &\approx\; g^{eq}(R) + \tfrac{1}{2}\,\frac{dg^{eq}(R)}{d\rho}\,\mathbf{R}\cdot\nabla\rho(\mathbf{r},t) \;\; , \end{aligned} \tag{55}$$

where $g^{eq}(R,t)$ is the equilibrium pair-correlation function for a homogeneous system with concentration $\rho(\mathbf{r},t)$, which is thus an implicit function of $\mathbf{r}$ and t, and the density derivative is with respect to $\rho(\mathbf{r},t)$. This implicit dependence is not denoted here explicitly for brevity. Substitution of the expansions (53,55) into the integral in eq.(52), using that $\nabla V(|\,\mathbf{r}-\mathbf{r}^{'}\,|) = -\hat{\mathbf{R}}\,dV(R)/dR$ (with $\hat{\mathbf{R}}=\mathbf{R}/R$ the unit vector along $\mathbf{R}$), together with,

$$\begin{aligned} \oint d\hat{\mathbf{R}}\;\hat{\mathbf{R}} \;&=\; \mathbf{0}\;, \\ \oint d\hat{\mathbf{R}}\;\hat{\mathbf{R}}\,\hat{\mathbf{R}} \;&=\; \frac{4\pi}{3}\,\hat{\mathbf{I}}\;, \end{aligned} \tag{56}$$

where the integrals range over all directions of $\mathbf{R}$, and $\hat{\mathbf{I}}$ is the identity tensor, leads to,

$$\rho(\mathbf{r},t)\int d\mathbf{r}'\; g(\mathbf{r},\mathbf{r}',t)\,\rho(\mathbf{r}',t)\,\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|) \;=\; \frac{2\pi}{3}\,\nabla\rho\,\frac{d}{d\rho}\left[\rho^2\int_0^\infty dR\;R^3\,\frac{dV(R)}{dR}\,g^{eq}(R)\right]. \tag{57}$$

We can now employ the standard equilibrium statistical mechanical expression for the pressure P^{eq},

$$P^{eq} \;=\; \rho\,k_BT - \frac{2\pi}{3}\,\rho^2\int_0^\infty dR\;R^3\,\frac{dV(R)}{dR}\,g^{eq}(R)\;, \tag{58}$$

to obtain Fick's diffusion equation (45) with,

$$D_c(\rho(\mathbf{r},t)) \;\equiv\; D_c^0\,\beta\,\frac{dP^{eq}(\rho(\mathbf{r},t))}{d\rho(\mathbf{r},t)}\;. \tag{59}$$

When gradients in concentrations are very large, such that the concentrations varies non-linearly on length scales set by the range R_V of the inter-molecular pair-potential, the Taylor expansions (53,55) are inaccurate, and higher orders of ∇ come into play in the diffusion equation.
Note that $(dP^{eq}/d\rho)\;\nabla\rho = \nabla P^{eq}$, which suggests that mass transport is driven by gradients in the pressure, which is intuitively appealing.
There is a subtlety for solutions of macromolecules. Instead of the mechanical pressure, the osmotic pressure appears in eq.(59) for the diffusion coefficient. This is a consequence of the fact that the macro-molecular pair-correlation function is thermally averaged with respect to the degrees of freedom of the solvent molecules. Furthermore, we neglected hydrodynamic interactions between such macromolecules, which do have an effect on the diffusive properties. A detailed discussion of these facts is beyond the scope of this chapter. The full expression for the diffusion coefficient of spherical macromolecules in solution, including the effects of hydrodynamic interactions, can be found in, for example, Ref.[12, 13, 14].

3.3 Diffusion of hydrogen through metals

Diffusion of hydrogen through metals has been studied since about 1850, and revived as a research area some decades ago due to the possible applications in energy storage. Hydrogen dissolves in metals not as intact H_2-molecules but as H-atoms, which reside at interstitial positions of the metal's crystal structure. There is a large number of neighbouring interstitial positions, so that hydrogen diffuses much faster as compared to diffusion of a molecule, where a thermal displacement requires the improbable event that there is a neighbouring vacancy (a missing atom in the crystal) to which it can move to. Fast diffusion of hydrogen is important for possible energy storage, since hydrogen must be dissolved in the metal (upon storage) and removed from the metal (upon use) within a reasonable time. Other important aspects for energy storage are the amount of hydrogen that can be dissolved, the effect of the mechanical properties of the metal on dissolving hydrogen, and the processes that occur right after absorption of hydrogen at the metal's surface, One of the metals in which large amounts of hydrogen can be dissolved is palladium. This is also one of the few metals that do not brittle and keeps to a large extent its original structural properties upon dissolving large amounts of hydrogen. We will therefore mainly focus on palladium. Right after absorption of hydrogen on the metal's surface, diffusion into the metal first requires the dissociation of H_2 into H-atoms, followed by surface diffusion of the atoms in order to find the appropriate locations to enter the metal crystal, after which diffusion into the bulk of the metal occurs. Dissolving hydrogen changes the lattice spacing of the crystal, mechanical stresses and crystal defects may be created, which both affect the diffusive properties of the H-atoms. Needless to say that a detailed discussion of all these complications can not be covered in this section. In the following we will briefly discuss the most important features of hydrogen-metal systems, and discuss diffusion from the gas phase into the metal on the basis of a simple diffusion model that accounts for crossing the surface of the metal. Besides transferring hydrogen into a metal by exposure to hydrogen gas, other methods can be used, like electrochemical deposition or by partially ionizing the gas phase, which circumvent the necessary dissociation after absorption to the metal's surface.

Let us first consider the phase behaviour of the hydrogen/palladium system. For low concentrations of hydrogen, it is homogeneously distributed within the palladium. For large concentrations, and sufficiently low temperatures, where the H-atoms strongly interact (through the crystal environment), phase separation is observed where a H-rich phase coexists with a H-poor phase. In both phases the hydrogen is disordered,while at room temperature the crystal lattice spacing for the H-poor phase is $0.3894\,nm$, and for the H-rich phase $0.4040\,nm$ [15, 16]. It seems likely that this phase equilibrium is similar to a gas-liquid phase coexistence, where "the gas" is the H-poor phase, and "the liquid" is the H-rich phase. The only difference between the two phases is their hydrogen concentration (and the difference lattice spacing of the palladium crystal), without any structural differences, like for a gas and a liquid. The experimental phase diagram is shown in Fig.8a [17, 18]. The line is the binodal, which has much the same form as for common gas-liquid coexistence. Notice that large amounts of hydrogen can be solved in palladium, where the ratio of hydrogen to palladium atoms is close to unity.

The simplest way to dissolve hydrogen in metals (palladium in particular), is by exposing the metal to gaseous hydrogen, of which the pressure can be varied by external means. Typically, more hydrogen dissolves when the hydrogen pressure is increased. First consider small pressures, such that H_2-molecules in the gaseous phase and H-atoms in the palladium environment do not interact with each other. For such ideal systems the chemical potential μ_g and μ_{Pd} of hydrogen in the gaseous phase and in palladium, respectively, depend on the corresponding

number concentrations ρ_g and ρ_{Pd} as,

$$\begin{aligned} \mu_g &= \mu_g^0(T) + k_B T \, \ln\{\Lambda^3 \, \rho_g\} \, , \\ \mu_{Pd} &= \mu_{Pd}^0(T) + k_B T \, \ln\{\Lambda^3 \, \rho_{Pd}\} \, , \end{aligned} \tag{60}$$

where $\mu_g^0(T)$ and $\mu_{Pd}^0(T)$ depend on temperature T only, and Λ is the Broglie wave length. Since an H_2-molecule splits in two H-atoms on dissolving in palladium, in equilibrium we have,

$$\mu_g \; = \; 2 \, \mu_{Pd} \, , \tag{61}$$

and hence,

$$\rho_{Pd} \; = \; f(T) \, \sqrt{p} \, , \tag{62}$$

where we used the ideal gas law $p = k_B T \, \rho_g$ for the pressure of the gaseous hydrogen phase, and where,

$$f(T) \; = \; \left[\, \frac{\Lambda^{-3}}{k_B T} \, \exp\left\{ \frac{\mu_g^0(T) - 2 \, \mu_{Pd}^0(T)}{k_B T} \right\} \right]^{1/2} \, , \tag{63}$$

is a function of temperature. That the amount of hydrogen that dissolves varies linearly with the square root of the pressure is commonly referred to as Sieverts's law [20]. Sieverts's law is verified experimentally, as can be seen in Fig.8b (for the $V_3 \, Ga \, H_x$-system). For higher pressures there are strong deviations from Sievert's law due to interactions between hydrogen molecules/atoms, as can be seen from Fig.8c [19]. Also note that more hydrogen dissolves at lower temperatures. Since diffusion typically slows down at lower temperatures, their are optimum operation conditions where a compromise must be found between appropriate time scales for hydrogen loading/recovery and the amount of hydrogen that can be stored.
Diffusion of hydrogen atoms within the bulk of the metal is described by Fick's law (46). The above mentioned changes in metal structure upon dissolving hydrogen can lead to a position dependent diffusion coefficient, or alternatively, to a contribution to the flux that relates to stresses that result from the crystal structure deformation. These effects will be neglected in the following. The diffusion of H_2-molecules in the gas phase is assumed to be sufficiently fast as compared to transfer rates to the metal, so that the concentration within the gas phase is a constant, independent of position. We will also assume that the hydrogen concentration in the gas is independent of time (which is the case when the gas container is very large compared to the volume of the metal, and/or when gas is conti nuously supplied). A complication that needs to be considered here is the above described processes that occur right after absorption of H_2-molecules onto the metal's surface. A detailed account of these surface processes is far beyond the scope of this chapter. Here, we simply lump all these processes in an energy barrier within a narrow region at the metal's surface that each molecule has to overcome on transfer from the gas phase to the metal bulk. The height of this barrier sets the ratio between the time-independent concentrations at the metal's surface on the gas-phase side and the metal-bulk side (see the sketch in Fig.9a). The surface-concentration ratio is proportional to $\exp\{-\beta \, E_b\}$, where E_b is the height of the energy barrier. The calculation of this energy is a complicated problem in itself, depending on all the complicated surface processes described above. We will refrain from considering these surface processes, and simply assume the concentration at the

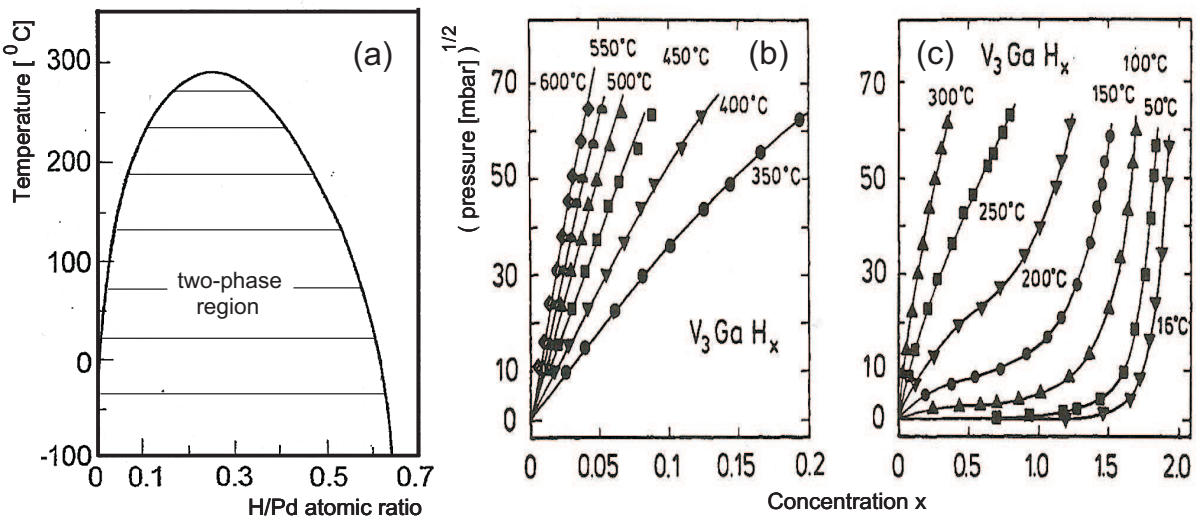


Fig. 8: *(a) The phase diagram of hydrogen/palladium. The solid line is the binodal corresponding to the coexistence between a H-rich and H-poor phase, of course both embedded within the palladium (taken from Ref.[17]). (b) The amount of dissolved hydrogen in the $V_3\,Ga\,H_x$-system for low pressures, and (c) for high pressures (taken from Ref.[19]).*

metal's surface, just inside the bulk, to be a given constant ρ_0^*. The diffusion process is thus described by the equations,

$$\begin{aligned} \frac{\partial}{\partial t}\,\rho(z,t) &= D\,\frac{\partial^2}{\partial z^2}\,\rho(z,t) \quad , \quad z>0\;,\; t>0\;, \\ \rho(z=0,t) &= \rho_0^* \quad , \quad t>0\;, \\ \rho(z,t=0) &= 0 \quad , \quad z>0\;. \end{aligned} \tag{64}$$

where z is the perpendicular distance from the metal's surface (as depicted in Fig.9a). The middle boundary condition states that the concentration at the surface is always equal to ρ_0^*, and the last initial condition ensures that the starting situation (at time $t=0$) is one where there is no hydrogen within the metal. The solution of this set of equations is,

$$\rho(z,t) \;=\; \rho_0^*\,\mathit{erfc}\left\{\frac{z}{\sqrt{4\,D\,t}}\right\}\;, \tag{65}$$

where "*the complementary error-function*" is defined as,

$$\mathit{erfc}\{x\} \equiv \frac{2}{\sqrt{\pi}}\int_x^\infty dw\,\exp\{-w^2\}\;. \tag{66}$$

These theoretical concentration profiles are plotted in Fig.9b. Obviously, the concentration profile spreads into the palladium as time evolves. In these plots, the value of the diffusion coefficient is taken equal to its typical value $D=10^{-7}\,m^2/s$. The total amount of hydrogen that diffused into the palladium at a certain time t is equal to,

$$\Delta\rho(t) \;=\; A\int_0^\infty dz\,\rho(z,t) \;=\; 4\,A\,\rho_0^*\,\sqrt{\frac{D\,t}{\pi}}\;, \tag{67}$$

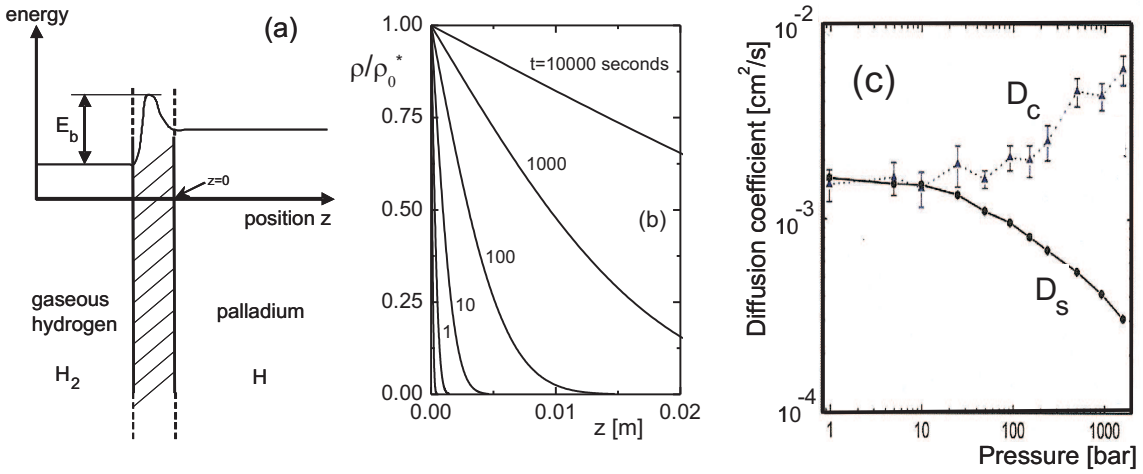


Fig. 9: *(a) A sketch of the simplified model for the kinetics of transfer of hydrogen from the gas phase to the metal bulk. (b) The concentration profiles (ρ/ρ_0^* versus z) for different times, according to eqs.(65,66) The value of the diffusion coefficient is taken equal to $D = 10^{-7}\ m^2/s$. (c) The self- and collective diffusion coefficients of hydrogen in $Zn(bdc)(ted)_{0.5}$, $bdc =$ benzenedicarboxylate, $ted =$ triethylenediamine (taken from Ref.[21]).*

where A is the area of the surface that is exposed to the hydrogen gas. The amount of stored hydrogen thus saturates as the square root of time. Typically, a palladium with a surface area of $A = 1\,m^2$ stores within an hour about $0.04 \times \rho_0^*[m^{-3}]$ hydrogen atoms. A reasonable estimate for ρ_0^* is a typical concentration in equilibrium, which is of the order $10^{27}\ H - atoms/m^3$. This gives an amount of hydrogen stored in one hour equal to about $100\,gram$. It is questionable whether this slow storage rate justifies an economic application, also in view of the high price of palladium [I am, however, hesitant to make a statement like this since my knowledge of economy is essentially equal to zero].

Both the self- and collective diffusion coefficients of hydrogen atoms are concentration dependent as a result of mutual interactions. This can be seen in Fig.9c, where both coefficients are plotted as a function the pressure of the gas phase for a certain metallic compound [21]. As explained in the introduction to section 3, the expected increase of the collective diffusion coefficient and the decrease of the self diffusion coefficient in case of repulsive $H - H$ interactions with increasing concentration is indeed observed. For larger concentrations of hydrogen and sufficiently low temperatures, however, attractions between H-atoms must be present since a gas-liquid phase separation is observed, as discussed above. Since gas-liquid phase transitions are abundant and occur in many different types of systems, the next section is devoted to the kinetics of gas-liquid phase separation from an initially unstable state, which is referred to as spinodal decomposition.

Solutions of Fick's diffusion equation for many types of geometries (like the infinite half-plane geometry in the above example) are discussed in Ref.[22].

4 A Negative Diffusion Coefficient: Spinodal Decomposition

As already mentioned at the end of the introduction to section 3, the collective diffusion coefficient can become negative when strong attractions between the molecules exist. This leads to a temporal increase of inhomogeneities (which are always present due to fluctuations) and ultimately leads to full phase separation. According to eq.(59) the collective diffusion coefficient is negative when $dP^{eq}/d\bar{\rho} < 0$ (with P^{eq} the equilibrium pressure of the homogeneous system and $\bar{\rho}$ the number density of the homogeneous system). The pressure thus decreases with increasing concentration, which is counter intuitive, and therefore hints to an instability. The kinetics of phase separation starting from an homogeneous, unstable system is referred to as **spinodal decomposition**. Phase separation from an initially meta-stable system, where nuclei are formed that grow in time, is referred to as **nucleation-and growth**. As will be seen in this section, the initial morphology in case of spinodal decomposition is qualitatively different from nucleation and growth. Instead of more-or-less ad-random distributed small nuclei of a significantly different concentration as compared to the initially homogeneous system, during spinodal decomposition a continuous growth of density differences is observed, where the relatively low and high concentration regions form a bi-continuous, space-spanning labyrinth structure.
The aim of this section is to quantitatively describe the temporal evolution of the density during the initial stages of phase separation of an unstable, initially homogeneous state. The point of departure is the generalized diffusion equation (52). This approach can be considered as the microscopic foundation of the classic **Cahn-Hilliard theory of spinodal decomposition**, which is based on irreversible thermodynamics [23, 24, 25].

4.1 An introduction to spinodal decomposition

Consider a homogeneous system that is unstable. In an experiment such a system may be prepared by suddenly cooling the system from a temperature in the stable region in the phase diagram to a temperature in the unstable part, where $dP^{eq}/d\bar{\rho} < 0$. Such a sudden change is commonly referred to as *a quench*. Macroscopic density inhomogeneities develop after the quench. The temporal evolution of the density is sketched in Fig.10. The inhomogeneous density can be thought of as being a superposition of sinusoidal density variations (in mathematical terms this refers to Fourier decomposition). A sinusoidally varying density is referred to as *a density wave*. As will be seen later in this section, during the initial stages of the phase separation, one of these density waves grows most fast. The wavelength corresponding to this most fast growing density wave is typically of a macroscopic size, of the order of hundreds of microns. In the *initial stage* of the phase separation, therefore, both the change $\delta\rho$ of the density and gradients of the density are small, as sketched in Fig.10 (upper panel). The initial stage is also referred to as the *linear regime*, since equations of motion for the density may be linearized with respect to $\delta\rho$. Then there is the so-called *intermediate stage*, where $\delta\rho$ is not small anymore, so that linearization is no longer allowed. Gradients of the density are still small, like in the initial stage, due to the long wavelengths that are unstable. This stage is depicted in Fig.10 in the second upper panel. Subsequently, the decomposition reaches the so-called *transition stage* where the lower and larger binodal concentrations ($\bar{\rho}_-$ and $\bar{\rho}_+$, respectively) are attained in various parts of the system, as sketched in Fig.10, the second lower panel. These binodal concentrations are the concentrations of the two phases that coexist when phase separation is completed. At this transition stage, sharp interfaces between the regions with concentrations close to $\bar{\rho}_-$ and $\bar{\rho}_+$ exist. Inhomogeneities are now large, and higher order terms in gradients

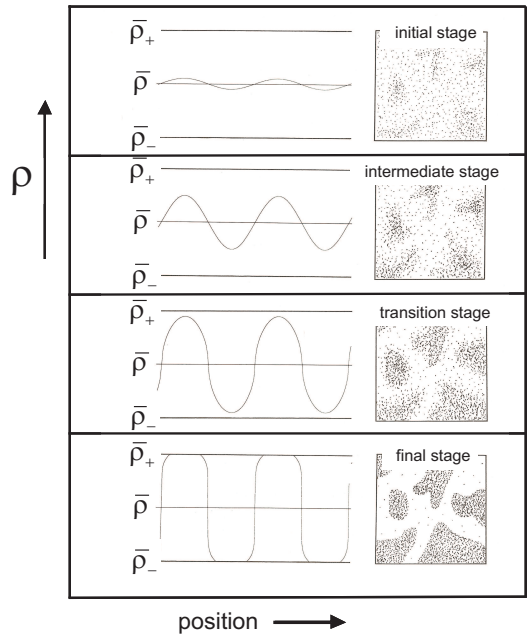


Fig. 10: *A sketch of the temporal development of the density after a quench into the unstable part of the phase diagram. Time increases from top to bottom. The left column of figures is a sketch of the density versus position, while the right column depicts a corresponding morphology of density variations. The concentrations $\bar{\rho}_+$ and $\bar{\rho}_-$ are the binodal concentrations, which are the concentrations of the two coexisting phases after phase separation is completed. Taken from Ref.[12].*

of the density come into play. In the *late stage* of the phase separation the interfaces develop : concentration gradients sharpen and the interfacial curvatures change to ultimately establish coexistence (see Fig.10, lower panel). We thus arrive at the following classification of the different stages during decomposition,

$$
\begin{aligned}
Initial\ stage \ &: \ \delta\rho/\bar{\rho}\ is\ small, \\
&\quad gradients\ are\ small\ (\text{"}diffuse\ interfaces\text{"}), \\
\\
Intermediate\ stage \ &: \ \delta\rho/\bar{\rho}\ is\ not\ small, \\
&\quad gradients\ are\ small\ (\text{"}diffuse\ interfaces\text{"}), \\
\\
Transition\ stage \ &: \ \delta\rho/\bar{\rho}\ is\ large, \\
&\quad gradients\ are\ not\ small\ (\text{"}sharp\ interfaces\text{"}), \\
\\
Final\ stage \ &: \ \delta\rho/\bar{\rho}\ is\ large, \\
&\quad gradients\ are\ large\ (\text{"}very\ sharp\ interfaces\text{"}).
\end{aligned}
$$

The terminology (i) "small", (ii) "not small" and (iii) "large" means that equations of motion for the density can be expanded (i) to leading order, (ii) to second order and (iii) all orders must be accounted for. Equations of motion for the density in the initial and intermediate stage can be expanded to leading order with respect to gradients of the density, while the leading non-linear contributions in $\delta\rho/\bar{\rho}$ must be included in the intermediate stage. The second order terms in

an expansion with respect to gradients of the density, which must be included in the transition stage, are referred to here as describing the dynamics of *sharp interfaces*, while all higher order terms must be included to realistically describe the dynamics of *very sharp interfaces* in the final stage. These very sharp interfaces have a width of the order of a few molecule diameters (except in case of quenches very close to the critical point, where the equilibrium interfacial thickness may be large).

4.2 Spinodal decomposition in the initial stage

In this section we describe the initial stage of demixing quantitatively, where the formation of inhomogeneities is a diffusive process. As before, let $\bar{\rho}$ denote the number density of colloidal particles in the homogeneous state, before decomposition occurred, and let $\rho(\mathbf{r},t)$ denote the macroscopic number density as a function of the position $\mathbf{r}$ in the system at time t after the system became unstable and started to demix. Define the change of the macroscopic density $\delta\rho(\mathbf{r},t)$ relative to that in the homogeneous state as,

$$\rho(\mathbf{r},t) \;=\; \bar{\rho} + \delta\rho(\mathbf{r},t)\;. \tag{68}$$

In the initial stage of the phase separation we have,

$$\frac{\mid \delta\rho(\mathbf{r},t)\mid}{\bar{\rho}} \;\ll\; 1\;, \tag{69}$$

allowing linearization of the generalized diffusion equation (52) with respect to the change $\delta\rho$ of the density.
Let δg denote the accompanied change of the pair-correlation function,

$$g(\mathbf{r},\mathbf{r}',t) \;=\; g_0(\mid \mathbf{r}-\mathbf{r}'\mid) \;+\; \delta g(\mathbf{r},\mathbf{r}',t)\;. \tag{70}$$

Here, g_0 is the pair-correlation function of the homogeneous system right after the quench, before phase separation occurred. To obtain a closed equation for $\delta\rho$, the change δg of the pair-correlation function must be expressed in terms of $\delta\rho$. Such a *closure relation* may be obtained as follows. An important feature is that the pair-correlation function in the integral in the diffusion equation (52) is multiplied by the pair-force $\nabla V(\mid \mathbf{r}-\mathbf{r}'\mid)$, so that a closure relation is only needed for small distances $\mid \mathbf{r}-\mathbf{r}'\mid \leq R_V$, with R_V the range of the pair-interaction potential. R_V is usually of the order of the size of the molecules. Relaxation of density variations over such small distances is much faster than the demixing rates of the very long unstable wavelengths, simply because it takes more time to displace colloidal particles over larger distances. On a coarsened time scale that is much larger than relaxation times of inhomogeneities that extend over distances of the order R_V, but which still resolves the phase separation process, the pair-correlation function in the integral in the generalized diffusion equation may be replaced by the equilibrium pair-correlation function. This is the statistical equivalent of the **thermodynamic local-equilibrium approximation**. The equilibrium pair-correlation function is to be evaluated at the instantaneous macroscopic density in between the positions $\mathbf{r}$ and $\mathbf{r}'$ (compare to what has been done in subsection 3.2). Hence, to first order in $\delta\rho$, and for $\mid \mathbf{r}-\mathbf{r}'\mid \leq R_V$,

$$\delta g(\mathbf{r},\mathbf{r}',t) = \delta g^{eq}(\mid \mathbf{r}-\mathbf{r}'\mid)\Big|_{density=\rho(\frac{\mathbf{r}+\mathbf{r}'}{2},t)} = \frac{dg^{eq}(\mid \mathbf{r}-\mathbf{r}'\mid)}{d\bar{\rho}}\,\delta\rho(\tfrac{\mathbf{r}+\mathbf{r}'}{2},t)\,, \tag{71}$$

and,

$$g_0(|\,\mathbf{r}-\mathbf{r}'\,|) \;=\; g^{eq}(|\,\mathbf{r}-\mathbf{r}'\,|)\,, \tag{72}$$

where g^{eq} is the equilibrium pair-correlation function for a homogeneous system with density $\bar{\rho}$ and the temperature after the quench. The two relations (71,72) are certainly wrong for distances $|\,\mathbf{r}-\mathbf{r}'\,|$ comparable to the wavelengths of the unstable density variations. For such distances the system is far out of equilibrium. The validity of the relations (71,72) is limited to small distances, where $|\,\mathbf{r}-\mathbf{r}'\,|\leq R_V$. Substitution of eqs.(71,72) into the generalized diffusion equation (52), renaming $\mathbf{R}=\mathbf{r}'-\mathbf{r}$, yields,

$$\begin{aligned}\frac{\partial}{\partial t}\,\delta\rho(\mathbf{r},t) \;&=\; D_c^0\,\Big\{\,\nabla^2\delta\rho(\mathbf{r},t)-\beta\,\bar{\rho}\,\nabla\cdot\int d\mathbf{R}\,[\nabla_R V(R)] \\ &\qquad\times\left(g^{eq}(R)\,\delta\rho(\mathbf{r}+\mathbf{R},t)+\bar{\rho}\,\frac{dg^{eq}(R)}{d\bar{\rho}}\,\delta\rho(\mathbf{r}+\tfrac{1}{2}\mathbf{R},t)\right)\Big\}\,,\end{aligned} \tag{73}$$

with ∇_R the gradient operator with respect to $\mathbf{R}$. The density can now be gradient-expanded like in subsection 3.2. We now need to extend the expansion to include the two higher order terms as compared to that in eq.(53), for reasons that will become clear in a moment. For example,

$$\begin{aligned}\rho(\mathbf{r}+\mathbf{R},t) \;&\approx\; \rho(\mathbf{r},t)+\mathbf{R}\cdot\nabla\rho(\mathbf{r},t) \\ &\qquad +\tfrac{1}{2}\,\mathbf{RR}:\nabla\nabla\rho(\mathbf{r},t)+\tfrac{1}{6}\,\mathbf{RRR}\,\vdots\,\nabla\nabla\nabla\rho(\mathbf{r},t)\,,\end{aligned} \tag{74}$$

where the vertical dots indicate contraction of adjacent indices (for example, $\mathbf{RR}:\nabla\nabla=\sum_{m,n=1}^{3}R_mR_n\nabla_n\nabla_m$). A similar expansion is made for $\rho(\mathbf{r}+\frac{1}{2}\mathbf{R},t)$. Substitution into eq.(73), noting that $[\nabla_R V(R)]=(\mathbf{R}/R)\;dV(R)/dR$, and integration with respect to the directions of $\mathbf{R}$, and using eqs.(56) together with,

$$\begin{aligned}\oint d\hat{\mathbf{R}}\;\hat{\mathbf{R}}\,\hat{\mathbf{R}}\,\hat{\mathbf{R}}\,\hat{\mathbf{R}} \;&=\; \mathbf{0}\,, \\ \oint d\hat{\mathbf{R}}\;\hat{\mathbf{R}}_i\,\hat{\mathbf{R}}_j\,\hat{\mathbf{R}}_m\,\hat{\mathbf{R}}_m \;&=\; \frac{4\pi}{15}\,[\,\delta_{ij}\,\delta_{mn}+\delta_{im}\,\delta_{jn}+\delta_{in}\,\delta_{mi}\,]\,,\end{aligned} \tag{75}$$

where δ_{ij} is the Kronecker delta, leads to, with some effort,

$$\frac{\partial\rho(\mathbf{r},t)}{\partial t} \;=\; D_c^0\left[\,\beta\,\frac{dP^{eq}}{d\bar{\rho}}\,\nabla^2\rho(\mathbf{r},t)+\Sigma\,\nabla^2\nabla^2\rho(\mathbf{r},t)\right]\,. \tag{76}$$

This equation of motion reproduces Fick's diffusion equation with the expression (59) for the collective diffusion coefficient, but with an additional contribution $\sim\nabla^2\nabla^2\rho$, with a proportionality factor equal to,

$$\Sigma \;=\; \frac{2\pi}{15}\,\bar{\rho}\int_0^\infty dR\,R^5\,\frac{dV(R)}{dR}\left(g^{eq}(R)+\frac{1}{8}\,\bar{\rho}\,\frac{dg^{eq}(R)}{d\bar{\rho}}\right)\,. \tag{77}$$

This constant is commonly referred to as **the Cahn-Hilliard square-gradient coefficient**. This higher order gradient contribution to Fick's diffusion equation is insignificant for stable systems, where $dP^{eq}/d\bar{\rho}>0$. For unstable systems where $dP^{eq}/d\bar{\rho}<0$, on the contrary, this contribution is essential, as will be seen shortly.

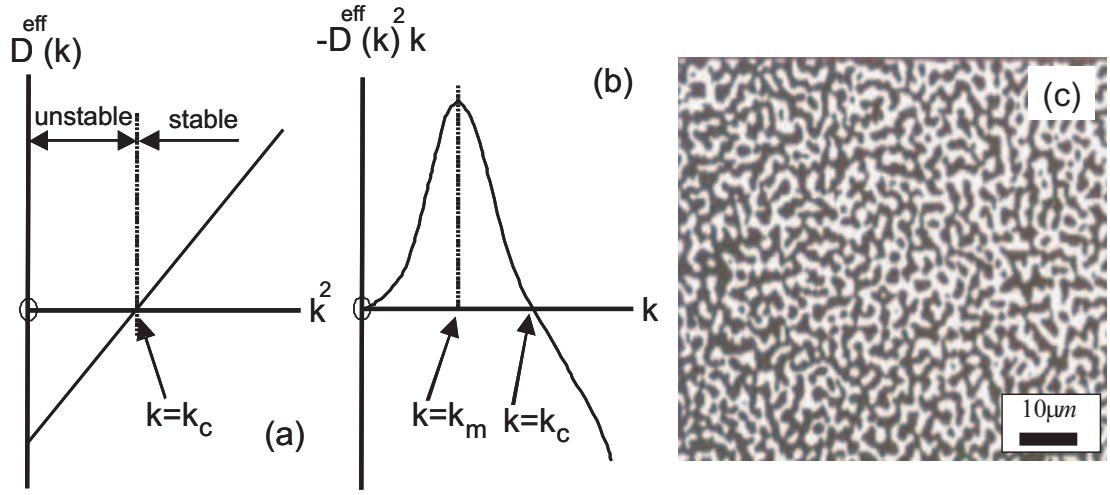


Fig. 11: *(a) The effective diffusion coefficient as a function of the wave vector. The critical wave vector k_c separates the unstable from the stable wave vectors. (b) The demixing rate as a function of the wave vector. k_m is the wave vector for which the corresponding density wave grows most fast. (c) The typical bi-continuous interconnected labyrinth structure of low (bright regions) and higher (dark regions) density that exists during the initial stages of spinodal demixing. This picture is taken from Ref.[26]*

An arbitrary spatially varying density can be decomposed in a sum of sinusoidal density variations with varying wave lengths. It thus suffices to consider the same initial sinusoidal density variation $\rho(\mathbf{r}, t=0) = \bar{\rho} + \delta\rho_0 \times \sin\{\mathbf{k}\cdot\mathbf{r}\}$ (as already introduced at the end of the introduction of section 3) where $\Lambda = 2\pi/k$ is the wavelength of the *density wave*, and the direction of the *wave vector* $\mathbf{k}$ defines the direction in which the density wave extends. Substitution of the Ansatz $\rho(\mathbf{r}, t) = \bar{\rho} + \delta\rho(t) \times \sin\{\mathbf{k}\cdot\mathbf{r}\}$ into eq.(76) immediately gives,

$$\delta\rho(t) \;=\; \delta\rho_0 \,\exp\left\{-D^{eff}(k)\,k^2\,t\,\right\}\;, \tag{78}$$

where the *effective diffusion coefficient* is equal to,

$$D^{eff}(k) \;=\; D_c^0\,\beta\left[\frac{dP^{eq}}{d\bar{\rho}} + \Sigma\,k^2\right]\;. \tag{79}$$

It follows from eq.(78) that the sinusoidal density variations for which $D^{eff}(k) < 0$ are unstable. For most systems $\Sigma > 0$, so that it follows from eq.(79) that there are unstable modes only when $dP^{eq}/d\bar{\rho} < 0$, which reproduces the classic thermodynamic instability criterion. Furthermore, only sufficiently small wave vectors are unstable and contribute to demixing. According to eq.(79) all wave vectors smaller than the critical wave vector,

$$k_c \;=\; \sqrt{-\frac{dP^{eq}/d\bar{\rho}}{\Sigma}}\;, \tag{80}$$

are unstable. That is, sinusoidal density variations with a wavelength larger than $\Lambda_c = 2\pi/k_c$ will grow in amplitude, giving rise to the creation of inhomogeneities. Density variations with a shorter wavelength remain stable (see Fig.11a). This is consistent with the assumption made in the above analysis that only long wavelengths (much larger than $2\pi/R_V$) are unstable.

Note that when the higher order spatial derivatives in the diffusion equation (76) (corresponding to the contribution $\sim \Sigma$ in eq.(79) for the effective diffusion coefficient) are omitted for unstable systems, arbitrary short wavelength density variations demix arbitrary fast. This is clearly unphysical, so that the higher order spatial derivatives are essential for a realistic description of demixing, contrary to the diffusion in stable systems.
The rate with which density variations grow, according to eq.(78), is equal to $-D^{eff}(k)\,k^2$. For,

$$k_m \;=\; \sqrt{-\frac{dP^{eq}/d\bar{\rho}}{2\Sigma}} \;=\; k_c/\sqrt{2}\;, \tag{81}$$

the growth rate attains a maximum value (see Fig.11b). The amplitude of the density variation with a wavelength equal to $\Lambda_m = 2\pi/k_m$ thus grows faster than any other demixing density wave. This introduces a characteristic length scale of inhomogeneities. Since the growth rates are independent of the direction of the wave vector, a bi-continuous interconnected labyrinth structure of low and higher density exists during the initial stages of demixing. A sketch of such a labyrinth structure is given in Fig.11c. The width of the labyrinth substructures are equal to the wavelength Λ_m of the most fast growing density wave.

4.3 The microscopic origin of the spinodal instability

To understand on a microscopic level why a system can become thermodynamically unstable, let us rewrite the generalized diffusion equation (52) as,

$$\frac{\partial}{\partial t}\,\delta\rho(\mathbf{r},t) \;=\; -\beta\,D_c^0\,\nabla\cdot\rho(\mathbf{r},t)\,\left[\mathbf{F}^{Br}(\mathbf{r},t)+\mathbf{F}^{Int}(\mathbf{r},t)\right]\;, \tag{82}$$

where,

$$\mathbf{F}^{Int}(\mathbf{r},t) \;=\; -\int d\mathbf{r}'\;g(\mathbf{r},\mathbf{r}',t)\,\rho(\mathbf{r}',t)\,\nabla V(|\,\mathbf{r}-\mathbf{r}'\,|)\;, \tag{83}$$

is the direct force, which stems from inter-molecular potential interactions (as specified by the pair-potential V) and,

$$\mathbf{F}^{Br}(\mathbf{r},t) \;=\; -k_BT\,\nabla\ln\{\rho(\mathbf{r},t)\}\;, \tag{84}$$

is the Brownian force on a molecule at the position $\mathbf{r}$. Now consider a molecule at $\mathbf{r}$ in an inhomogeneous environment, as sketched in Fig.12a. The inhomogeneous macroscopic density may be thought of as an instantaneous realization of the fluctuating density. A little thought shows that the Brownian force is always directed towards the region with lower concentration, as depicted in Fig.12a. Now suppose that the pair-interaction potential is purely attractive. The direct force $\mathbf{F}^{Int}$ is then directed in the opposite direction, towards the region with a larger density, since in that region there are more neighbouring molecules attracting the molecule under consideration : this can also be seen formally from eq.(83) for the direct force, using a purely attractive pair-interaction potential. On lowering the temperature, the Brownian force diminishes, since that force is directly proportional to the temperature. The direct force, however, increases in magnitude, due to the fact that the pair-correlation function becomes more pronounced (to leading order in the density this follows from the expression $g = \exp\{-V/k_BT\}$, where $V<0$ for an attractive pair-potential). At the temperature where $|\,\mathbf{F}^{Int}\,| > |\,\mathbf{F}^{Br}\,|$, the net force on the molecule is directed towards the region with a larger density. This is the mechanism that is responsible for **uphill diffusion**, and leads to phase separation.

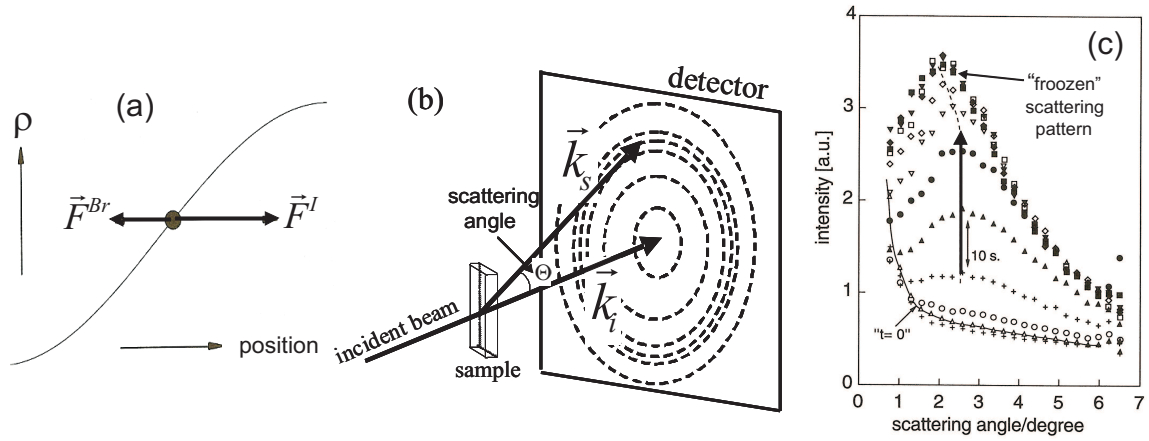


Fig. 12: *(a) The direct and Brownian force on a molecule, indicated by •, in an inhomogeneous system. For an attractive direct force* $\mathbf{F}^I$*, its direction is towards the region with larger concentration, as sketched here. (b) A schematic of the scattering set up. The wave vector is the equal to* $\mathbf{k}_s - \mathbf{k}_i$*, with* $\mathbf{k}_s$ *the scattered wave vector and* $\mathbf{k}_i$ *the incident wave vector, which both have a length equal to* $2\pi/\lambda$*, where* λ *is the wavelength of the light beam. (c) The scattered intensity from a suspension of stearyl-coated silica particles as a function of the scattering angle. Each curve is measured* 10 *seconds after each other. At later times, the scattering peak "freezes", due to the formation of a gel in the regions of higher concentration. From Ref.[27].*

4.4 A light scattering experiment

The above predictions can be verified by means of time resolved light scattering. It can be shown that the time-dependent scattered intensity $I(t)$ at very small scattering angles from an inhomogeneous system is given by (see, for example, Ref.[12]),

$$I(k,t) \ \sim <|\,\delta\rho(\mathbf{k},t)\,|^2> \ , \tag{85}$$

where the brackets $< \cdots >$ refer to ensemble averaging over initial conditions. The wave vector can be varied through variation of the scattering angle Θ (the angle between the incident beam and the direction of detection: see Fig.12b)). In case of scattering, the wave vector is equal to $\mathbf{k} = \mathbf{k}_s - \mathbf{k}_i$, where $\mathbf{k}_i$ and $\mathbf{k}_s$ point in the direction of the incident beam and the direction of detection, respectively. In case of elastic scattering, both the incident and scattered wave vector $\mathbf{k}_s$ and $\mathbf{k}_i$ have the same length equal to $2\pi/\lambda$, where λ is the wavelength of the light. It follows that $k = \frac{4\pi}{\lambda}\sin\{\Theta/2\} \approx \frac{2\pi}{\lambda}\,\Theta$, for small scattering angles Θ. The wave vector thus determines the position on a screen at which scattered light is collected, as sketched in Fig.12b. Substitution of eq.(78) into eq.(85) gives,

$$I(k,t) \ \sim < \delta\rho_0^2 > \ \exp\left\{-2\,D^{eff}(k)\,k^2\,t\right\} \ . \tag{86}$$

This result predicts that in the initial stage of decomposition, $\ln\{I(k,t)\}$ is a linear function in time and that the intensity develops a "ring-like scattering pattern", where the position of the peak grows at a time-independent position. This is indeed observed in some systems, like in dispersions of stearyl silica colloidal particles [27]. As can be seen from the plots in Fig.12c, the ring-like scattering pattern for this particular system "freezes" at some time during demixing. This is due to the formation of a gel-like phase in the more concentrated region of the labyrinth structure. In other systems (see, for example, Refs.[28, 29]), the peak position of the scattering

ring is observed to shift to lower scattering angles right from the start of the experiment. Such a shift of position of the peak of the ring-like scattering pattern is expected in the intermediate stage of demixing (the above analysis has been extended to include the intermediate stage in Refs.[12, 30]). Probably the time range within which the linear theory applies is already past at the time a first reliable measurement could be performed.
Spinodal decomposition is also found for hydrogen dissolved in palladium, within the two-phase region in Fig.8a. The required attractive interactions between the H-atoms is probably due to palladium-crystal-environment mediated interactions.

References

[1] R. Festa, E. Galleani d'Agliano, Physica A **90**, 229 (1978)

[2] L. Gunther, M. Revzen, A. Ron, Physica A **95**, 367 (1979)

[3] M.P. Lettinga, E. Grelet, Phys. Rev. Lett. **99**, 197802 (2007)

[4] E. Pouget, E. Grelet. M.P. Lettinga, Phys. Rev. E **84**, 041704 (2011)

[5] C. Dalle-Ferrier, M. Krüger, R.D.L. Hanes, S. Walta, M.C. Jenkins, S.U. Egelhaaf, Soft Matter **7**, 2064 (2011)

[6] K. Kang, J.K.G. Dhont, J. Chem. Phys. **122**, 044905 (2005)

[7] K. Kang, J.K.G. Dhont, J. Chem. Phys. **124**, 044907 (2006)

[8] K. Kang, J.K.G. Dhont, J. Chem. Phys. **126**, 214501 (2007)

[9] P.-G. de Gennes, J. Chem. Phys. **55**, 572 (1971)

[10] K.E. Evans, S.F. Edwards, J. Chem. Soc., Faraday Trans. **2**, 1891 (1981)

[11] D. Richter, M. Monkenbusch, A. Arbe, J. Colmenero, Adv. Polym. Sci. **174**, 1-221 (2005)

[12] J.K.G. Dhont, *An Introduction to Dynamics of Colloids*, Elsevier, Amsterdam, 1996

[13] G. Nägele, Physics Reports **272**, 215 (1996)

[14] G. Nägele, J.K.G. Dhont, G. Meier, *Diffusion in Condensed Matter*, Chapter 16, page 619, (eds. P. Heitjans, J. Kärger), Springer Verlag, Berlin.

[15] A.J. Maeland, Canadian Journal of Physics **46**, 121 (1968)

[16] T.B. Flanagan, W.A. Oates, Ann. Rev. Mater. Sci. **21**, 269 (1991)

[17] J. Vökl, G. Alefeld, *Diffusion in Solids: Recent Developments*, (eds. A.S. Nowick, J.J. Burton), Academic press, New York,1975

[18] I.S. Anderson, D.K. Ross, C. Carlile, Phys. Lett. **68**A, 249 (1978)

[19] M. Schlereth,H. Wipf, J. Phys.: Condens. Matter **2**, 6929 (1990)

[20] A. Sieverts, Z. Phys. Chemie **88**, 451 (1914)

[21] J. Liu, J.Y. Lee, L. Pan, R.T. Obermyer, S. Simizu, B. Zande, J. Li, S.G. Sankar, J.K. Johnson, J. Phys. Chem. C **112**, 2911 (2008)

[22] J. Crank, *The Mathematics of Diffusion*, Oxford Science Publications, Clarendon Press, Oxford, 1975

[23] J.W. Cahn, J.E. Hilliard, J. Chem. Phys. **28**, 258 (1958), **31**, 688 (1959)

[24] M. Hillert, Acta Metallica **9**, 525 (1961)

[25] J.W. Cahn, J. Chem. Phys. **42**, 93 (1965)

[26] I. Demyanchuk, S.A. Wieczorek, R. Holyst, J. Chem. Phys. **121**, 1141 (2004)

[27] H. Verduin, J.K.G. Dhont, J. Coll. Interf. Sci. **172**, 425 (1995)

[28] F. Mallamace, N. Micali, S. Trusso, S.H. Chen, Phys. Rev. E **51**, 5818 (1995)

[29] P. Wiltzius, F.S. Bates, W.R. Heffner, Phys. Rev. Lett. **60**, 1538 (1988)

[30] J.K.G. Dhont, J. Chem. Phys. **105**, 5112 (1996)

B 4 Elementary Excitations in Crystalline Matter: Phonons and Magnons [1]

K. Schmalzl
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Any kind of matter can be seen as a system of interacting particles, where the structure is determined by its ground state and the dynamics is represented by excited states. Ground states and excited states are determined by the Coulomb interaction between ions and electrons. Due to this interaction the excited states are collective excitations.
Dynamical properties of matter are directly related to physical properties, which on their part can be measured in experiments. To understand phenomena like temperature dependent effects, superconductivity, phase transition or also magnetic properties, to name just a few, the basic elementary excitations have to be understood.
In this chapter two types of elementary excitations in matter will be discussed. First the concept of lattice waves in the harmonic approximation will be introduced. An outlook to anharmonic effects which lead to temperature dependent phenomena, and a short comparison to excitations in amorphous media will then be given.
Analogous to the periodic sequence of atoms in a lattice, the periodic arrangement of electronic spins of atoms leads to spin waves. On the basis of a model Hamiltonian the examples of ferro- and antiferromagnetism will be treated. The quasiparticle of the spin waves are the magnons, similar to the phonons for the lattice waves.
The main part of the formalism presented here deals with an infinite periodic system. As interactions in matter can be manifold and complicated appropriate approximations will be presented.

2 Lattice Excitations: Phonons

2.1 Adiabatic Approximation

A solid represents a quantum mechanical many-body system consisting of interacting nuclei and electrons. The Hamiltonian contains the kinetic and potential energy with all nuclear and electronic coordinates and momenta and their interaction.

$$H = H_{\mathrm{nucl}} + H_{\mathrm{el}} + H_{\mathrm{el-nucl}}.$$

The wavefunction is thus a function of all nuclear, electronic, and spin coordinates. As this is a rather complicated system in total, approximations are needed: With the adiabatic approximation (also called Born-Oppenheimer approximation) one can decouple the electron from the nuclear dynamics: One uses the fact that the light electrons adjust instantaneously to the slower motion of the heavy ionic cores (with the electron mass being about 10^{-4} of the ion mass!). During the motion of the atomic cores the electrons follow the ions adiabatically and remain in their ground state, the energy of which, however, is given by the ionic configuration.

2.2 Harmonic Approximation

The ions move in a potential V, made up of the potential energy of the ion configuration and of the energy of the electrons in this configuration.
Assuming small atomic displacements $\mathbf{u}_l$ out of their equilibrium positions, the potential energy V can be written as a Taylor series in these atomic displacements, and the Hamiltonian reads

$$H_{\mathrm{ion}} = \sum_l \frac{\mathbf{p}_l^2}{2m_l} + V_0 + \sum_{l,\alpha} \frac{\partial V}{\partial u_{l,\alpha}} u_{l,\alpha} + \frac{1}{2} \sum_{l,\alpha} \sum_{l',\beta} \frac{\partial^2 V}{\partial u_{l,\alpha} \partial u_{l',\beta}} u_{l,\alpha} u_{l',\beta} + \dots . \qquad (1)$$

If one neglects all terms higher than of second order, indicated by dots, this approximation is called the harmonic approximation. (This approximation is made for convenience: the harmonic equations of motions are easier to solve and this approximation is justified as the harmonic term is the dominant one and gives good results at least at low temperatures.)
A closer look at equation (1) shows: V_0 is a constant and not relevant for the dynamics. The first-order term vanishes since it is assumed that the particles vibrate around their equilibrium positions:

$$H_{\text{ion}} = \sum_l \frac{\mathbf{p}_l^2}{2m_l} + \frac{1}{2}\sum_{l,\alpha}\sum_{l',\beta} \frac{\partial^2 V}{\partial u_{l,\alpha}\partial u_{l',\beta}} u_{l,\alpha}u_{l'\,\beta}. \tag{2}$$

The second-order derivatives of the potential are the force constants

$$\Phi_{\alpha\beta}(ll') = \frac{\partial^2 V}{\partial u_{l,\alpha}\partial u_{l',\beta}}\,. \tag{3}$$

They describe the force acting on the atom l in the direction α when the atom l' is displaced in the direction β.
In the harmonic approximation, the atomic vibrations can be transformed into decoupled normal vibrations (as we see later, leading for periodic systems to (harmonic) phonon dispersion curves). As these latter vibrations have no coupling, they have infinite lifetime. For lifetime (linewidth) or frequency shift and other temperature effects one would have to consider the higher-order terms neglected in the harmonic approximation. They are called the anharmonic terms.
The considerations above apply to crystalline as well as amorphous solids.

2.3 Normal Coordinates and Phonons

To finally investigate vibrations of particles in our system we have to solve the equations of motion,

$$m\frac{\partial^2 u_{l,\alpha}}{\partial t^2} = -\frac{\partial V}{\partial u_{l,\alpha}} = -\sum_{l',\beta}\Phi_{\alpha\beta}(ll')u_{l',\beta}.$$

Because of the harmonic approximation we can formulate a harmonic ansatz for the time dependence of the displacements $\mathbf{u}$ for all l,

$$\mathbf{u}_l \propto \exp(-i\omega t) \tag{4}$$

and get as result

$$\sum_{l',\beta}\left(\Phi_{\alpha\beta}(ll') - \omega^2 m_l\delta_{l,l'}\delta_{\alpha,\beta}\right)u_{l',\beta} = 0\,. \tag{5}$$

The sum runs over all atoms in the system. With N atoms, this is a $3N \times 3N$-dimensional problem resulting in a $3N \times 3N$ force-constant matrix.
After diagonalizing this matrix, the diagonal elements are the squared frequencies. All eigenfrequencies ω_λ are represented by eigenvectors $\chi(l,\alpha|\lambda)$ with quantum numbers λ, and the general solution is

$$u_{l,\alpha}(t) = \sum_\lambda Q_\lambda \chi(l,\alpha|\lambda)e^{-i\omega_\lambda t}\,. \tag{6}$$

The superposition coefficients Q_λ are the so-called normal coordinates.

Upon quantization, the normal coordinates can be expressed in terms of phonon annihilation and creation operators, $a_\lambda, a_\lambda^\dagger$,

$$Q_\lambda \rightarrow \sqrt{\hbar/2\omega_\lambda}\left(a_\lambda + a_\lambda^\dagger\right). \tag{7}$$

With the eigensolutions we can also formulate the Hamiltonian (1) in terms of these solutions or normal coordinates. In this representation the Hamiltonian will then be a set of uncoupled harmonic oscillators, where each of them corresponds to a collective mode. This decoupling is possible only within the harmonic approximation,

$$H = \sum_\lambda \hbar\omega_\lambda \left(n_\lambda + \tfrac{1}{2}\right) \tag{8}$$

with $\frac{1}{2}\hbar\omega$ being the lowest energy state of a quantum harmonic oscillator and

$$n_\lambda = a_\lambda^\dagger a_\lambda \tag{9}$$

the occupation number operator with the commutation relations for creation and annihilation operators, $[a_\lambda, a_{\lambda'}^\dagger] = \delta_{\lambda,\lambda'}, [a_\lambda, a_{\lambda'}] = 0 = [a_\lambda^\dagger, a_{\lambda'}^\dagger]$; $\hbar\omega_\lambda$ is the excitation energy quantum of the collective mode, the phonon energy.

2.4 Bloch's Theorem and Vibrations in Crystals

A crystal is an infinite periodic system, and with Bloch's theorem (or mathematically Floquet's theorem) we can make use of this lattice periodicity. The theorem states that the displacements of same atoms in different unit cells differ only by a phase factor. We denote now not only the elementary cell by an index l but also the atom within a cell by an index κ. Then the displacements u can be written as

$$\mathbf{u}(l\kappa) = \frac{1}{\sqrt{m_\kappa}}\mathbf{e}(\kappa|\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{R}_l}. \tag{10}$$

(The mass factor is for later convenience.) The first part describes the periodicity of the cells, the second one is the wave-like part.
Now the equation of motion takes the form

$$\omega^2(\mathbf{q})e_\alpha(\kappa|\mathbf{q}) = \sum_{\kappa',\beta} D_{\alpha\beta}(\kappa\kappa'|\mathbf{q})e_\beta(\kappa'|\mathbf{q}). \tag{11}$$

This equation now treats only atoms within one unit cell. The so-called dynamical matrix $D_{\kappa\kappa'}(\mathbf{q})$ is the Fourier transform of the force constant matrix (3), multiplied by the square root of the masses of the atoms (in order for (11) to obtain the normal form of an eigenvalue equation),

$$D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \frac{1}{\sqrt{m_\kappa m_{\kappa'}}}\sum_{l'} \Phi_{\alpha\beta}(l,\kappa,l',\kappa')e^{i\mathbf{q}\cdot(\mathbf{R}_{l'}-\mathbf{R}_l)} \tag{12}$$

(D-definition after Maradudin et al. [2]).
If r is the number of atoms in the elementary cell, now the $3N \times 3N$ problem of (5) is reduced to a $3r \times 3r$ problem in (11). For each wavevector $\mathbf{q}$ the secular problem

$$\left|D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) - \omega^2(\mathbf{q})\delta_{\alpha,\beta}\delta_{\kappa,\kappa'}\right| = 0 \tag{13}$$

has to be solved. The diagonalization of the dynamical matrix leads to the eigenfrequencies $\omega_j(\mathbf{q})$ with the branch index j. The quantum number λ for general systems above is now detailed by two quantum numbers $\mathbf{q}$ and j for lattices. For each eigenfrequency $\omega_j(\mathbf{q}) = \omega_j(-\mathbf{q})$ with $j = 1, \ldots, 3r$ a corresponding eigenvector $\mathbf{e}(\kappa|\mathbf{q}, j)$ to equation (11) exists.

As we have seen we can gain all phonon properties from considering only one unit cell. In the same way as a crystal is composed of a periodic array of unit cells also the phonon properties show a periodicity:

$$\omega_j(\mathbf{q}) = \omega_j(\mathbf{q} + \tau),$$
$$\mathbf{e}(\kappa|\mathbf{q}, j) = \mathbf{e}(\kappa|\mathbf{q} + \tau, j)\,.$$

The vector τ is a so-called reciprocal-lattice vector.
But still in an experiment one would not necessarily find the same result in different Brillouin zones (different τ): the scattering function contains terms involving

$$(\mathbf{q} + \tau) \cdot \mathbf{e}(\kappa|\mathbf{q}, j) e^{i(\mathbf{q}+\tau)\cdot\mathbf{R}(\kappa)},$$

which might lead to different scattering intensities in different Brillouin zones for the same frequency!
Because of the hermiticity of the matrix the squared frequencies $\omega_j^2(\mathbf{q})$ must be real, and with this $\omega_j(\mathbf{q})$ can be either real or imaginary. In the latter case, for a calculated imaginary frequency, the crystal shows an instability. The curves $\omega_j(\mathbf{q})$ as a function of q for a given direction of $\mathbf{q}$ are called dispersion curves (analogous to the band-structure curves in the electronic case). The solutions of equation (11) or (13) describe collective modes for which all ions in a lattice move with the same time dependence but with a phase shift with respect to each other. The eigenvectors form a complete set of solutions that can be used as basis for any arbitrary motion of the ions.
With r atoms in the unit cell we will find $3r$ phonon branches per $\mathbf{q}$-vector. The factor 3 describes thereby the three spatial degrees of freedom of each atom within the elementary cell. Common to the phonon dispersion of all solids are the branches with lowest frequencies, which are the acoustic modes which start from the center of the Brillouin zone with a linear dependence on q. A simple picture of acoustic modes can be drawn with one of the simplest models, the monatomic linear chain.

2.5 Phonon Dispersion Curves

Consider a linear chain with atoms, all with mass M, separated by the unit cell length a. And assume an average force constant C acting between neighboring planes. This is the model of a monatomic linear chain. Let U_i be the displacement of plane i, With the periodic time dependence $U_i(t) = u_i e^{-i\omega t}$ the equations of motion (5) reduce to

$$-M\omega^2 u_i = C(u_{i+1} + u_{i-1} - 2u_i)\,.$$

The solution of these equations are plane waves $u_{i\pm1} = u_i e^{\pm iqa}$ periodic with wavevector q. For the dispersion relation $\omega(q)$ of the monatomic linear chain we get:

$$\omega(q) = (4C/M)^{1/2} |\sin\frac{qa}{2}|.$$

This dispersion is schematically shown in fig. 1. For one atom in one dimension we will get only one phonon branch. The sine dependence shows us: For small q-vectors (long wavelengths) the

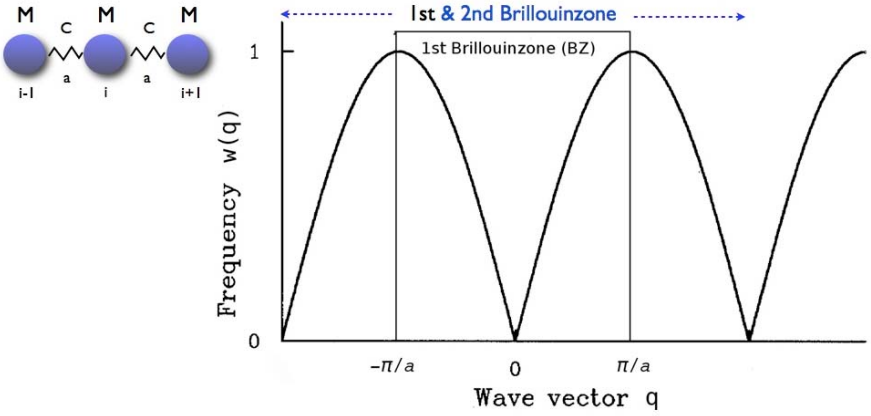


Fig. 1: *Schematic representation of the dispersion curve of the monatomic linear chain. The phonon dispersion can be limited to the first Brillouin zone.*

frequency ω is linear in q, $\omega = c_s q$ with the velocity of sound c_s. Also we see: for small q-vectors $q \rightarrow 0$ the frequency tends to zero, $\omega \rightarrow 0$.
One striking feature is the periodicity of the function: the period is $2\pi/a$, which is equal to a unit cell length in reciprocal space. We have seen already that one can add a reciprocal lattice vector $\tau = 2\pi/a$ to any point at any wavevector, the frequency and equations of motion will be identical to the first one. All information is contained within $-\pi/a < q < \pi/a$. This range is called the first Brillouin zone. Further zones are called the second, third, etc. Brillouin zone. One can limit the phonon dispersion to the first Brillouin zone.
The values $q = \pm\frac{\pi}{a}$ are the Brillouin zone boundaries lying half way between centers. In our case the slope of the dispersion is zero at the zone boundary, $d\omega/dq = 0$ for $q = \pm\frac{\pi}{a}$.

Phonon modes are grouped in acoustic and optic. Of the $3r$ possible modes, three modes are acoustic, and the remaining $(3r-3)$ modes are optic with a a non-zero frequency at $\mathbf{q} = 0$.
In the acoustic modes in the long-wavelength limit $q \rightarrow 0$ all atoms move with nearly the same phase. The wavelength λ of this movement is much larger than the lattice constant a. Depending

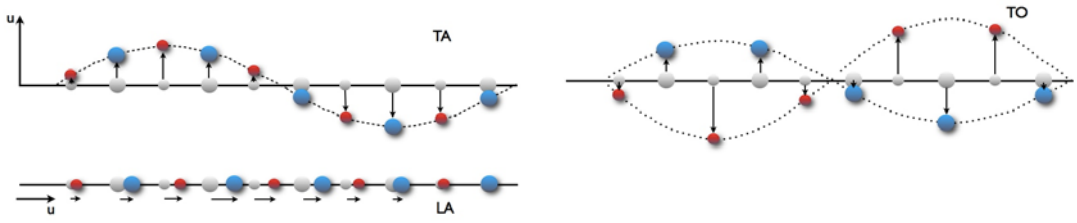


Fig. 2: *Schematic displacement patterns for a two-atomic linear chain for* $\mathbf{q} \neq 0$. *Top left:transverse acoustic mode (TA), bottom left: longitudinal acoustic mode (LA). Right: transverse optic mode (TO).*

on the orientation of the polarization vectors $\mathbf{e}(\kappa|\mathbf{q}j)$ with respect to $\mathbf{q}$, the modes can possibly be further classified to be longitudinal or transverse. In the longitudinal mode the displacement is parallel to the $\mathbf{q}$-vector, whereas in the two transverse modes the displacement is perpendicular. This may be true, however, only in main symmetry directions, e.g., for cubic symmetries along [100], [110], or [111] (edge, face, or space diagonal). In more complicated structures or

away from main symmetry directions the modes show a mixture of these displacement patterns and cannot be classified as purely longitudinal or transverse movements anymore.
Phonon dispersion curves $\omega_j(\mathbf{q})$ are usually plotted along high-symmetry directions where degeneracies may occur. They are obtained experimentally, e.g., from inelastic scattering of neutrons, photons, or atoms, and theoretically, e.g., from model or ab initio calculations. For crystals of the same symmetry group similar dispersion curves are obtained, the details depending upon the specific atom masses and force constants. Scales for frequencies are typically $10^{12}\,\mathrm{s}^{-1}$ (THz) $\equiv 33.36\,\mathrm{cm}^{-1} \equiv 4.136\,\mathrm{meV}$.

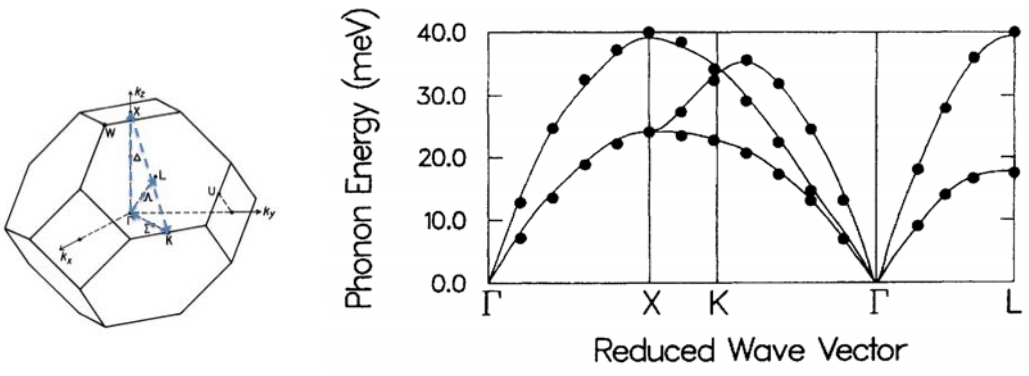


Fig. 3: *Phonon dispersion for aluminium from first principle ab initio methods in comparison with data from inelastic neutron scattering. [3] Left: Brillouin zone of the fcc lattice with the measured high symmetry directions marked.*

Figure 3 shows the phonon dispersion curve for the example of aluminium along the main symmetry directions of the fcc Brillouin zone. With one atom in the primitive unit cell the phonon dispersion shows three acoustic branches (one LA, two TA). The two lower transverse acoustic branches are degenerate in $\Gamma - X$ and $\Gamma - L$-direction which is characteristic of the symmetry group (Fm3m) of fcc aluminium. Also can be seen that the modes show zero curvature at the zone boundaries.
With two and more atoms in the unit cell the dispersion will show also optic modes. In ionic crystals these modes may be dipole-active, depending on the involved atoms, and can then be excited optically, e.g., by infra-red light (therefore they are called optic modes).
Figure 4 shows an example for a phonon dispersion with three atoms per unit cell, CaF_2, along the three main symmetry directions of the fcc Brillouin zone and along several other directions along the Brillouin zone faces. Depicted are also the group-symmetrical representations of the different phonon branches.
In ionic crystals where the different atoms are (differently) charged a macroscopic longitudinal electric field is introduced with the displacement of the atoms. As this field is of longitudinal character, longitudinal optic modes at $q = 0$ couple to this field and their energy is increased, whereas transverse modes are not affected. This fact is described by the Lyddane-Sachs-Teller relation for ionic crystals for long wavelengths:

$$\frac{\epsilon(0)}{\epsilon_\infty} = \frac{\omega_L^2}{\omega_T^2}.$$

This means, in ionic crystals the degeneracy is partly removed at the Brillouin zone center for

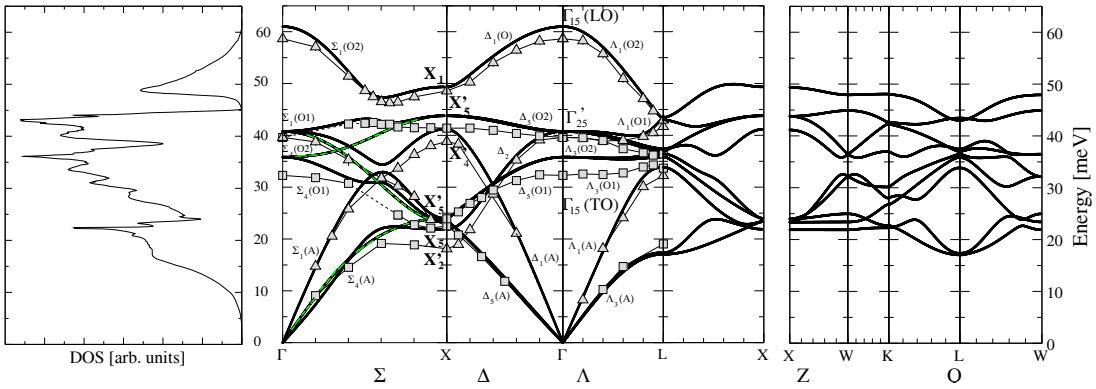


Fig. 4: *DOS and phonon dispersion curves for CaF_2 from inelastic neutron scattering (symbols) and ab initio method. [4] Triangles mean phonons measured in longitudinal, squares in transverse polarization.*

optic modes. The LO modes have higher frequencies than the TO modes. This LO–TO splitting might be anisotropic and direction-dependent.
For identical atoms in the unit cell, the modes at the Brillouin zone center are degenerate. With increasing charge difference the gap between the LO and TO modes increases.

2.6 Phonon Density of States

A convenient way to display a phonon spectrum is also the phonon density of states (DOS). The phonon spectrum with a a given dispersion $\omega_j(\mathbf{q})$ can be characterized by the one-phonon density of states, $D(\omega) = \sum_{\mathbf{q},j} \delta(\omega - \omega_j(\mathbf{q}))$. It counts the number of phonon modes within a frequency interval at a given frequency ω by summing over all branches in the first Brillouin zone and is the sum of all frequencies in all directions in reciprocal space (not only those in a chosen dispersion direction). It is shown on the left hand side of fig. 4. Pronounced peaks can be seen coming from flat parts of the branches.

2.7 Anharmonic Effects and Phase Transitions

Within the harmonic approximation we introduced before many experimental results can be reproduced but temperature dependent effects cannot be predicted. In experiment and real life those effects are omnipresent. To model and understand effects like thermal expansion, effects coming from phonon-phonon collision or phase transitions, etc., the higher order terms in the Hamiltonian (1) have to be taken into account. With this the decoupling of vibrations is not possible anymore and an interaction of phonons takes place.
In experiment temperature effects will be seen in a shift of the phonon frequency compared to the harmonic one, in thermal expansion and a phonon linewidth broadening. The linewidth is the inverse of the lifetime of a phonon and can be seen via a damping of the spectral lines in experiment. This is ascribed to decay processes where a phonon decays after a certain time into other phonon states.
The thermal expansion is the process easiest to consider. With increasing temperature in general the volume of the material increases. One can expect that frequencies decrease with increasing volume as bond lengths increase and interatomic forces decrease.

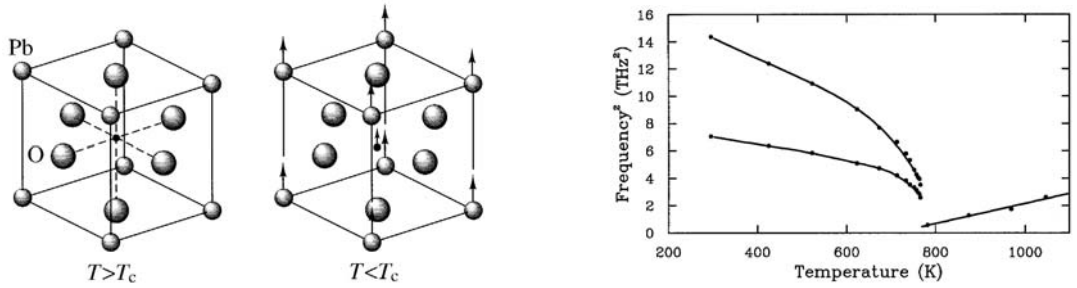


Fig. 5: *$PbTiO_3$ undergoes a cubic-tetragonal phase transition at 763 K. Left: distortion associated with the transition involves small atomic displacements in [001]. Right: Decreasing frequency when approaching the transition temperature. Below the phase transition a splitting in two modes occurs. Taken from [5])*

The change of volume and frequency with changing temperature can be described with the mode-Grüneisen-parameter:

$$\gamma_{\mathbf{q},j} = -\frac{\partial \ln \omega(\mathbf{q}, j)}{\partial(\ln V)}.$$

For the shift of phonon frequencies decay and scattering processes have to be taken into account, including higher order terms than the harmonic ones as well as thermal expansion, all depending on higher order terms in the potential.
When calculating the phonon frequencies of the structure known to exist at high temperatures with the harmonic equations of motion (11), the squared phonon frequencies ω^2 can be negative, leading to imaginary frequencies. This means the assumed structure is not stable at 0 K and an another structure with lower energy exists. The structural distortion can be identified when finding the eigenvector corresponding to the imaginary frequency.
An example for the change of frequency with temperature is the second order displacive phase transition. From high to low temperatures the system seems to undergo small structural modifications. Figure 5 shows the temperature variation of the two-fold degenerate TO frequency near the ferroelectric phase transition in $PbTiO_3$. When cooling in an experiment from high to low temperatures the phonon frequency will decrease from the stable harmonic value at high temperatures to zero at the transition. At this point the crystal is unstable against the corresponding distortion and undergoes a phase transition to a lower energy phase. This mode which can be found in the high temperature phase is called a soft mode. Below the phase transition there will be another mode increasing in frequency when cooling. Below T_c a splitting into two modes occurs due to the reduced symmetry.

2.8 Other Thermal Effects

The internal energy of a phonon system is

$$U(T) = \sum_{\mathbf{q},j} \hbar\omega_j(\mathbf{q}\left(\langle n_{\mathbf{q},j}\rangle + \tfrac{1}{2}\right) \tag{14}$$

with the Bose occupation number

$$\langle n_{\mathbf{q},j}\rangle = \frac{1}{1 - e^{\hbar\omega_j(\mathbf{q})/k_B T}}. \tag{15}$$

There are two effects from temperature: The one comes from the explicit T-dependence of the Bose factor; the other comes from the (anharmonic) T-dependence of $\omega_j(\mathbf{q})$. Neglecting the latter is called the quasi-harmonic approximation.
Similarly to the internal energy the entropy, free energy, specific heat, etc. can be calculated.

2.9 Amorphous Media and Liquids

So far we considered perfect lattice symmetries.
If only single atoms of the crystal are replaced by defects, then still an average lattice periodicity on a long-range scale is preserved. The wavevectors are still approximately good quantum numbers, and the notions of Brillouin zone etc. are good approximations. In contrast, in disordered materials like amorphous materials, glasses or liquids there is less or no long-range order at all leading to structural disorder. As a consequence, the $\mathbf{q}$-vector is not well defined and no longer a good quantum number.

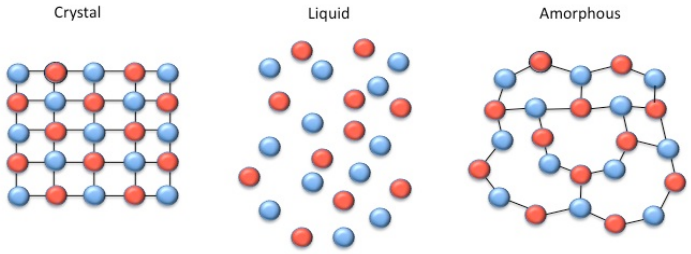


Fig. 6: *Schematic presentation of a crystalline, liquid and amorphous ordering. No long-range order or large distance periodicity exists in the latter two.*

But there exists still short range order and the local environment of a given atom is, similar to the crystalline case with nearest and next-nearest neighbor distances, relevant. With increasing distance from a given atom, these structural quantities are less well defined and become increasingly smeared out. With this loss of periodicity also no well defined Brillouin zones can be identified, one speaks about pseudo-Brillouin zones, where the size of the zones is defined by the next neighbor distance.
Nevertheless, in the long-wavelength limit glasses and disordered materials behave like crystals with propagation of sound waves. In this limit the phonon dispersion is similar to the one found in a crystalline system. With increasing wavelength vibrational properties lose their plane-wave character, which can be seen as a broadening of spectral distribution in experiment coming from the disorder of the system. This means for small q a linear dispersion can be found which is the macroscopic velocity of sound in the system and corresponds to the longitudinal acoustic phonon branch.
Figure 7 shows measured longitudinal dispersion curves for different glass systems. The modes show a pseudo-periodic behavior, where the extent of the Brillouin zone is determined by the nearest neighbor distance. Depending on the system, the minimum in the second Brillouin zone, which would correspond to the next Brillouin zone center, might be more pronounced like it is in the case of ethanol hinting to an order at slightly larger length scales. Liquids and amorphous materials are different from crystalline systems in that the former have strongly reduced transverse forces. Anyway, for reduced symmetries the terms longitudinal and transverse might lose their meaning.

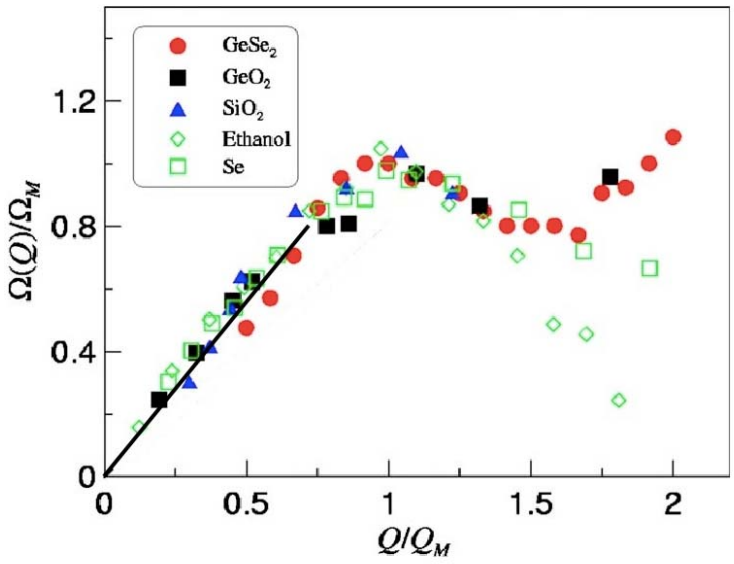


Fig. 7: *Dispersion curves of various glass-like systems measured with inelastic neutron scattering, rescaled for comparison in respect to frequency and wavevector Q [6]. The straight line indicates the velocity of sound.*

As a result of structural disorder, e.g., the heat conduction will be smaller in glasses and liquids than in crystalline materials due to defect-induced scattering. In a simple model, the heat conduction is given by

$$\kappa = \frac{1}{3} c_V v l$$

with the specific heat c_V, velocity v and mean free path length l. The structural or compositional disorder reduces the average mean free path and the thermal conductivity becomes reduced in comparison with that of a perfect crystal lattice.

3 Spin Waves: Magnons

So far we considered lattice vibrations, where the position of the atoms changes periodically leading to phonons as quantized particles. We will focus now on properties coming from the electron spin, which have been neglected so far. An electron system with its quantum-mechanical exchange interaction will lead to a spin-ordered ground state and its elementary excitations, the spin waves. The elementary excitations of a periodic spin system are of a wavelike character in analogy to lattice vibrations, and, similarly to the phonons, magnons as quantized particles of the spin wave, can be introduced. In this chapter the equation for spin wave excitations in a ferromagnet as well as in an antiferromagnet will be deduced. Usually magnetic order exists below a critical temperature at which a phase transition takes place. The temperature and magnetic field dependence will be treated shortly at the end of this chapter.
For the spin ordered ground state one can imagine different realizations, like all spins parallel (ferromagnetism), or antiparallel (antiferromagnetism), (anti-) ferrimagnetic ordering or more complex like helix-form.

Here we will focus on the most prominent cases of a spin ordering with parallel and antiparallel spins, leading to ferro- and antiferromagnetism.

$$\text{FM}: \ldots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \ldots \qquad \text{AFM}: \ldots \uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \ldots$$

Fig. 8: *Classical ferromagnetic (FM) and antiferromagnetic (AFM) ground state in a Heisenberg chain. The spin states S_i^z=+S and S_i^z=-S are pictured by '↑' and '↓'. The AFM ordering is described by two sublattices with opposite magnetization.*

The exchange interaction between magnetic moments leads to the alignment of the spins. The energy of the interaction between two atoms i and j and their spins $\mathbf{S}_i$ and $\mathbf{S}_j$ can be formulated within the Heisenberg model:

$$U = -2 \sum_{ij}{}' J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j,$$

where J_{ij} is the exchange integral describing the exchange energy of overlapping charge distributions of the different atoms i and j. $\sum'$ means summation over $i \neq j$. For a ferromagnetic system the exchange integral is defined as positive, $J_{ij} > 0$. For antiferromagnetics J is taken as negative for up and down spins, $J_{ij} < 0$.

3.1 Spin Waves in Ferromagnets

Ferromagnetic materials show a spontaneous magnetic moment also without an external magnetic field. In the ground state all (total) atomic spins are aligned parallel, and the magnetic and chemical unit cells are identical.
The z-component of a single spin S is $m_S = S$. Exciting this state, the z-component becomes reduced, $m_S < S$. In an energetically more favorable excitation state the excitations are distributed over all spins and the orientation of the spin vector relative to the lattice changes periodically forming a spin wave. The quantized particles of this spin wave are magnons.
In the Heisenberg-model the Hamiltonian of spins in an external magnetic field $\mathbf{B}$ can be written in the following way:

$$H = -\sum_{ij}{}' J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + \gamma \mu_B \mathbf{B} \cdot \sum_i \mathbf{S}_i. \tag{16}$$

The exchange integral J_{ij} decays fast with increasing distance between the atoms i and j, and in practice the interaction is limited to next neighbors. The second term is the Zeeman energy $-\mu_e \cdot \mathbf{B}$ which presents the potential energy of the spin magnetic moments with an external magnetic field $\mathbf{B}$; one typically chooses $\mathbf{B} = -B(0, 0, 1)$. The ions form a regular lattice, each

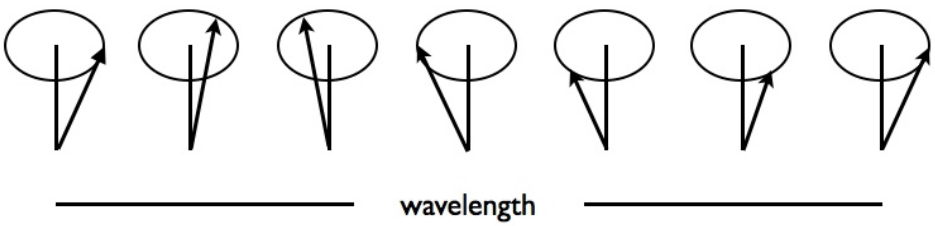


Fig. 9: *Schematic representation of a spin wave with wavelength λ.*

site is associated with a magnetic moment μ, with $\mu_e = -\gamma\mu_B \mathbf{S}_i/\hbar$, where γ is the Landé factor or gyromagnetic ratio ($\gamma \approx 2.00232$ for a single electron), μ_B is the Bohr magneton.

In general the Hamiltonian can contain further contributions, coming, e.g., from crystalline anisotropies coupling the spins to certain crystallographic directions, or the dipole-dipole interaction of the spins, etc. These contributions are usually smaller than the terms written down and shall be neglected.

We are now not dealing anymore with displacements (or creation and annihilation of displacement states) as in the phonon case but with creation and annihilation of magnon states. To describe excitations or the change in the z-component we introduce spin operators:

$$S_+ |m\rangle \propto |m+1\rangle\,, \quad S_- |m\rangle \propto |m-1\rangle\,, \quad S_z |m\rangle = m\,|m\rangle \tag{17}$$

Note that $\mathbf{S}\cdot\mathbf{S} = \frac{1}{2}(S_+S_- + S_-S_+) + S_z^2$. The spin operators fulfill the commutation relations: $[S_+, S_-] = 2S_z, [S_z, S_\pm] = \pm S_\pm$.

To determine the eigenstates of (16) we relate the spin operators to Bose creation and annihilation operators $a^\dagger$ and a. This is the Holstein-Primakoff transformation. The spin operators can then be written as:

$$\begin{aligned} S_+^i &= S_x^i + iS_y^i = \sqrt{2S}\left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2} \cdot a_i, \\ S_-^i &= S_x^i - iS_y^i = \sqrt{2S}a_i^\dagger\left(1 - \frac{a_i^\dagger a_i}{2S}\right)^{1/2} = (S_+^i)^\dagger, \\ S_z^i &= S - a_i^\dagger a_i. \end{aligned} \tag{18}$$

The last equation describes the deviation from the maximum value S by the number operator $a_i^\dagger a_i$, which will turn out to be the quantized excitations. The commutation relations for the creation and annihilation operators are the same as they were in the phonon case.

As in the phonon case, we assume a periodic lattice. The Hamiltonian (16) is invariant against a translation of basis vectors within the lattice. This means, solutions of the equation of motion in plane wave form will exist, and we will make a Bloch ansatz as for lattice waves, equation (10). It is then convenient to let the creator and annihilators act on collective states given by the wavevector $\mathbf{q}$:

$$a_i = \frac{1}{\sqrt{N}}\sum_{\mathbf{q}} b_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}^i} \text{ and } a_i^\dagger = \frac{1}{\sqrt{N}}\sum_{\mathbf{q}} b_{\mathbf{q}}^\dagger e^{i\mathbf{q}\cdot\mathbf{R}^i}. \tag{19}$$

The commutation relations for a_i and $a_i^\dagger$ are also valid for the $b_{\mathbf{q}}$ and $b_{\mathbf{q}}^\dagger$.

In the following we will consider low excitations with $S - m_S \ll S$. Similar to the harmonic approximation for the phonons where we expanded the potential only up to second order we will now only consider the lowest terms in a_i or $b_{\mathbf{q}}$. With this the square root in equation (18) is

approximately $\left(1-\frac{a_i^\dagger a_i}{2S}\right)^{1/2} \approx 1$, and the spin operators result to:

$$\begin{aligned} S_+^i &= \sqrt{\frac{2S}{N}} \sum_{\mathbf{q}} b_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{R}^i}, \\ S_-^i &= \sqrt{\frac{2S}{N}} \sum_{\mathbf{q}} b_{\mathbf{q}}^\dagger e^{i\mathbf{q}\cdot\mathbf{R}^i}, \\ S_z^i &= S - \frac{1}{N} \sum_{\mathbf{q}\mathbf{q}'} b_{\mathbf{q}}^\dagger b_{\mathbf{q}'} e^{i(\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}^i}. \end{aligned} \tag{20}$$

These equations are of a similar type as the displacements in a lattice in the harmonic approximation. The applied approximation assumes small deviations of the spins from their original direction and thus only small excitation energies. This would be the case for small temperatures but the approximation has proven to be reasonable in many cases.
The solutions are thus collective wave-like states, which are the spin waves. In an analogy to the phonons, the quantized particles of these vibrations are called (ferromagnetic) magnons which are created and destroyed by the operators $b_{\mathbf{q}}^\dagger, b_{\mathbf{q}}$.
With the ansatz (20) in the Hamiltonian (16) and with the restriction to a Bravais lattice where $J_{ij} = J_{0,i-j}$ one assumes low excitations and truncates the expansion after second-order terms in $b_{\mathbf{q}}^\dagger, b_{\mathbf{q}}$. The Hamiltonian for spin waves in ferromagnets results to

$$H = E_0 + \hbar\omega(\mathbf{q}) n_{\mathbf{q}} \quad \text{with} \quad n_{\mathbf{q}} = b_{\mathbf{q}}^\dagger b_{\mathbf{q}}. \tag{21}$$

with the ground-state energy depending on the interaction of the spins with an external magnetic field

$$E_0(B) = -NS^2 {\sum_{\mathbf{h}}}' J_{0\mathbf{h}} - \gamma\mu_B BNS \tag{22}$$

and the excitation energy

$$\hbar\omega(\mathbf{q}) = 4S \sum_{\mathbf{h}} J_{0\mathbf{h}} \sin^2(\mathbf{q}\cdot\mathbf{R}^{\mathbf{h}}/2) + \gamma\mu_B B \approx S \sum_{\mathbf{h}} J_{0\mathbf{h}} (\mathbf{q}\cdot\mathbf{R}^{\mathbf{h}})^2 + \gamma\mu_B B. \tag{23}$$

The result is valid with the assumption of small values of $\langle n_{\mathbf{q}} \rangle$ and neglects magnon-magnon interaction.
Equation (23) shows that the dispersion $\omega(\mathbf{q})$ is quadratic in q for small q, i.e. for long wavelengths. Unlike the case of phonons where $\omega(\mathbf{q})$ is linear in q for long wavelengths. The magnon dispersion is in the same energy range as the phonon dispersion, namely meV, see fig. 10.
Similar to a phonon dispersion the magnon dispersion can have several branches, e.g., different atoms in the magnetic unit cell or different sublattices can lead to not only one acoustic branch but additionally to one or several optic branches.

3.2 Spin Waves in Antiferromagnets

A simple model of an antiferromagnet assumes a lattice with two different sublattices with opposite spin directions like shown in fig. 3, where adjacent atoms belong to different sublattices. The total magnetization results to zero. An example is the lattice of the antiferromagnetic insulator MnF_2 crystal.

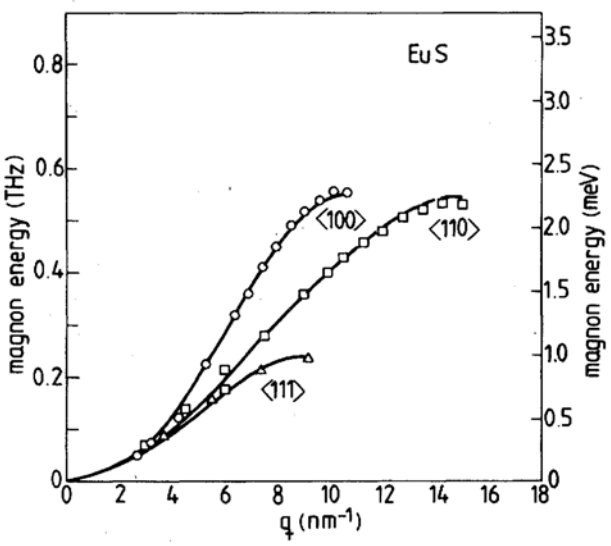


Fig. 10: *Magnon dispersion for ferromagnetic EuS from inelastic neutron scattering showing a strong anisotropy for different directions in* $\mathbf{q}$ *space. [7]*

The Hamiltonian has to include now the two different spin types, and the exchange integral is $J_{ij} < 0$ for an antiferromagnet. Like in the ferromagnet case we limit the interaction to next neighbors:

$$H = \sum_{ij} J_{ij} \mathbf{S}_1^i \cdot \mathbf{S}_2^j - \gamma \mu_B B \sum_i (\mathbf{S}_{1z}^i - \mathbf{S}_{2z}^j). \tag{24}$$

For simplicity, the Landé g-factor is assumed to be the same.
The following transformations are similar to the ferromagnet case, only one sort of spin (1) is aligned parallel to the z-direction, the other one antiparallel (2), as can be seen in the z-component of the spin operator. We employ again the Holstein-Primakoff transformation to replace the spin operators by boson operators. Now we have to take into account the two spin operators for the different sublattices. The equation equivalent to (18) is:

$$\begin{aligned} S_{1+}^i &= \sqrt{2S}\left(1 - \frac{a_{1i}^\dagger a_{1i}}{2S}\right)^{1/2} \cdot a_{1i}, \\ S_{2+}^i &= \sqrt{2S} a_{2i}^\dagger \left(1 - \frac{a_{2i}^\dagger a_{2i}}{2S}\right)^{1/2}, \\ S_{1z}^i &= S - a_{1i}^\dagger a_{1i}, \quad S_{2z}^i = -S + a_{2i}^\dagger a_{2i} \end{aligned} \tag{25}$$

and $S_{1-}^i = (S_{1+}^i)^\dagger$, $S_{2-}^i = (S_{2+}^i)^\dagger$.
We perform the same transformation to operators in $\mathbf{q}$-space, equation (19), and the truncation of the expansion in $b_{\mathbf{q}}$,

$$H = E_0 + 2S \sum_{\mathbf{h}} J_{0,\mathbf{h}} \sum_{\mathbf{q}} \sum_{\sigma=1,2} b_{\mathbf{q}\sigma}^\dagger b_{\mathbf{q}\sigma} + 2S \sum_{\mathbf{q}} \sum_{\mathbf{h}} J_{0,\mathbf{h}} e^{i\mathbf{q}\cdot\mathbf{R}^\mathbf{h}} (b_{\mathbf{q}1} b_{\mathbf{q}2} + b_{\mathbf{q}1}^\dagger b_{\mathbf{q}2}^\dagger). \tag{26}$$

The middle term counts elementary excitations on each sublattice. The last term contains cross-terms describing a coupling between two different sublattices. To finally diagonalize the Hamil-

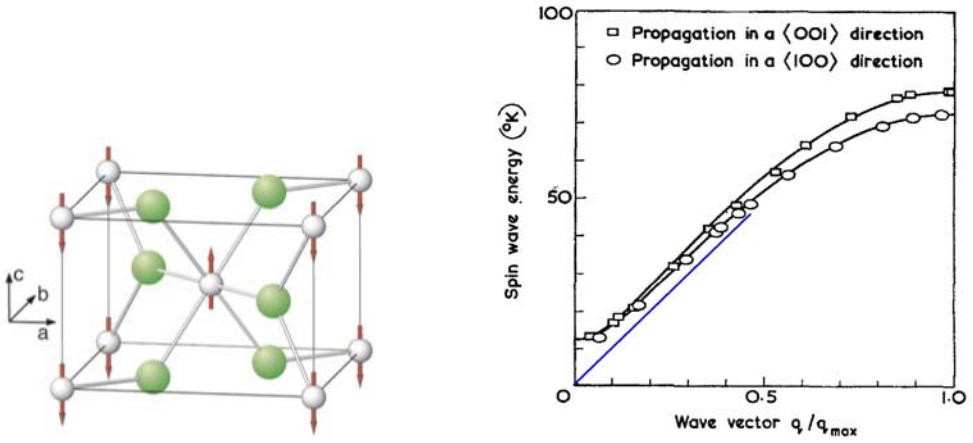


Fig. 11: *Left: Crystal and magnetic structure of MnF_2. The grey spheres represent Mn^{2+} ions, the arrows their relative direction of spin. [8] Right: Magnon dispersion for two different directions in k space. [9]*

tonian this coupling has to be removed via a transformation called the Bogoliubov transformation:

$$b_{\mathbf{q}1} = u_{\mathbf{q}} c_{\mathbf{q}1} + v_{\mathbf{q}} c^{\dagger}_{\mathbf{q}2}, , \quad b_{\mathbf{q}2} = v_{\mathbf{q}} c^{\dagger}_{\mathbf{q}1} + u_{\mathbf{q}} c_{\mathbf{q}2}.$$

Here the magnon creators of one sublattice are related to annihilators in the other sublattice where the size of each part might vary with the substance under consideration.
The requirement that the new operators $c^{\dagger}$ and c be Bose creation and annihilation operators lead to $u^2_{\mathbf{q}} - v^2_{\mathbf{q}} = 1$. The resulting Hamiltonian contains terms with $c^{\dagger}c$ (as one expects), but also terms with two creation or annihilation operators, $c^{\dagger}c^{\dagger}$ and cc as in the last term in (26). The second requirement is the vanishing of the latter. This fixes the u and v and the final form of the Hamiltonian.
Finally the Hamiltonian for antiferromagnetic spin waves results to:

$$H = E_0 + \sum_{\mathbf{q}\sigma} \hbar\omega_{\mathbf{q}} \left(c^{\dagger}_{\mathbf{q}\sigma} c_{\mathbf{q}\sigma} + 1/2 \right). \tag{27}$$

Different from ferromagnetic magnons the Hamiltonian contains here the number operators of two elementary excitations with the same energy and a zero point contribution.
With the approximation for small q and $B = 0$:

$$\hbar\omega_{\mathbf{q}} \approx 2S \left(\sum_{\mathbf{h}} J_{0,\mathbf{h}} (\mathbf{q} \cdot \mathbf{R}^{\mathbf{h}})^2 \right)^{1/2}. \tag{28}$$

We see that in an antiferromagnet the dispersion $\omega(\mathbf{q})$ for small q or long wavelength is linear in q as it was for the phonon dispersion. As an example this linear dispersion measured by inelastic neutron scattering is shown in fig. 11. In this special case the dispersion does not follow the linear dispersion in q but exhibits a gap at $q = 0$. This is due to an anisotropy field, which removes the degeneracy of the two different magnons leading to a finite energy at $q = 0$.

3.3 Temperature Dependence of the Magnetization

Magnons can be thermally excited similar to phonons. Without external magnetic field and at very low temperatures we assume the magnetization $M(0) = g\mu_B NS$ saturated and all spins

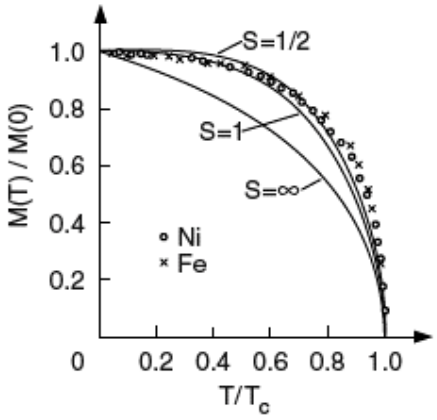


Fig. 12: *Reduced magnetization vs. reduced temperature for ferromagnetic Ni and Fe. The solid lines are calculations with different values for the total spin. [1]*

aligned. With increasing temperature the spontaneous magnetization $M(T)$ will decrease. We have seen before that the deviation from the saturated magnetization $M(0) \propto S$ is described by the occupation number operator $n_{\vec{q}}$.
In a ferromagnet the Curie temperature T_C separates the not-ordered paramagnetic phase, where the spontaneous magnetization vanishes, from the ordered ferromagnetic phase at $T < T_C$. The transition temperature for antiferromagnets is the Néel temperature T_N.
For temperatures below the Curie temperatures T_C the magnetization $M(T)$ for ferromagnets reads:

$$M(T) = \gamma\mu_B\Big(NS - \sum_{\mathbf{q}}\langle n_{\mathbf{q}}\rangle\Big), \tag{29}$$

where the brackets, $\langle \dots \rangle$, denote the thermal average. The saturation magnetization of one sublattice in an antiferromagnet is smaller than NS, and not all spins are parallel but deviate from a perfect antiparallel spin alignment. The diminution depends e.g., on the type of lattice or structure. The term in brackets, is the total spin in the z-direction. For a cubic lattice with lattice constant a and small temperatures $T \ll T_C$ the relative deviation of the magnetization from its saturated value results to:

$$\frac{M(0) - M(T)}{M(0)} = \frac{2.315}{S}\frac{V}{4\pi^2 a^3}\left(\frac{k_B T}{2SJ}\right)^{3/2}. \tag{30}$$

This $T^{3/2}$-law of the spontaneous magnetization is also found experimentally, see fig. 12.
For an antiferromagnet in analogy to (29) the z-component of the magnetization of the first sublattice results to

$$M_1 = \gamma\mu_B\Big(NS - \sum_{\vec{k}}\langle b^\dagger_{\mathbf{q}1} b_{\mathbf{q}1}\rangle\Big) \tag{31}$$

The temperature dependence of the magnetization results in approximation for a cubic lattice to

$$\frac{M_1(0) - M_1(T)}{\gamma\mu_B} = \frac{V}{4\pi^2 a^3\sqrt{2z}}\left(\frac{k_B T}{2JS}\right)^2. \tag{32}$$

(z is the number of neighbors.) Unlike in a ferromagnet the magnetization of an antiferromagnetic sublattice decreases with T^2 for small temperatures. The different temperature behavior is related to the different q-dependence of the dispersion curves at long wavelengths.

3.4 Dependence on magnetic field

So far we stated that in a ferromagnet at low temperatures well below T_C in general all magnetic moments are aligned parallel. Still one might find that the magnetization of the system increases in an applied magnetic field until reaching a saturation value. This is ascribed to the existence

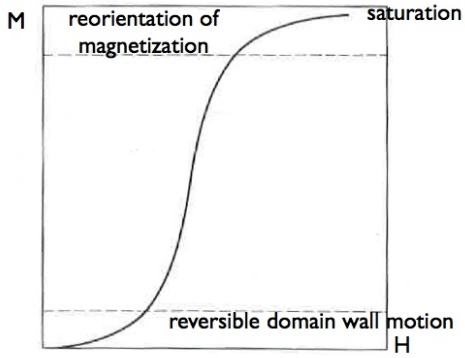


Fig. 13: *Schematic magnetization curve, magnetization vs. magnetic field.*

of domains, first postulated by P. Weiss (1907) and well established by now. In each of these domains the magnetization is saturated but oriented differently in respect to each other leading to lower total moment. The domain configuration might then change with an applied magnetic field. Within one domain however, one can expect to find well aligned spins, as the exchange interaction J_{ij} is short-range and limited to almost next-neighbors.
With applying an external magnetic field first the volume of the domains which are oriented already in direction of the external field will grow, then in a stronger external field, all domains will be oriented in direction of the external field.

4 Conclusion

As a conclusion we have seen here the formalism for elementary excitations for lattice and spin waves or respectively their quasiparticles phonons and magnons.
We saw first that we can decouple the electron from the ion dynamics and that it's enough to treat the ion movement in a potential V. The second major approximation was the truncation of this potential after the second order term. This allowed a harmonic ansatz for the displacements leading to a set of uncoupled harmonic oscillators / collective modes. Assuming an infinite periodic crystal the equations could be simplified and it was sufficient to look only at the dynamics within one unit cell or the first Brillouin zone. From the equations of motions the dynamical matrix and frequencies and eigenvectors and finally the dispersion curves resulted. For small q the acoustic-mode dispersion is linear in q approaching the velocity of sound. The phonon spectrum can also be presented with the phonon density of states . For temperature dependent effects higher order terms in the potential expansion have to be taken into account.
In case of amorphous materials, where the lattice periodicity is more or less missing, the short range order still leads to a phonon dispersion at small $\mathbf{q}$ vectors.

We treated the spin waves in a similar formalism as the lattice waves. Within the Heisenberg-model we determined the eigenstates with relating the spin operators to the creation and annihilation operators (Holstein-Primakoff transformation). Only that instead of creating or annihilat-

ing displacement states we were dealing now with magnon states. For an infinite periodic lattice we could make a Bloch ansatz for plane waves as for the phonons. And similar to the harmonic approximation where we expanded the potential only up to second order we considered only lowest terms in the creation and annihilation operators. In ferromagnets the spin wave dispersion curves were found to be quadratic in q, in antiferromagnets linear in q. The formalism for antiferromagnets differs from the one for ferromagnets as one has to consider two different spin-sublattices. This calls for another transformation called the Bogoliubov transformation, necessary to decouple the sublattices and be able to diagonalize the Hamiltonian.
In general the magnetization vanishes with increasing temperature. The transition temperature is called Curie- and Néel temperature for ferro- and antiferromagnets. The formalism presented was based in general on the assumption of having only one magnetic domain. In reality many domains might exist in a sample, leading to an increase of magnetization in external magnetic field.

References

[1] Textbooks on Solid State Physics:

W. Ludwig, *Festkörperphysik II* (Akademische Verlagsgesellschaft Frankfurt am Main 1970)
U. Rössler, *Solid State Theory: An Introduction* (Physica Verlag, Springer, Berlin Heidelberg 2009)

[2] A. A. Maradudin, E. W. Montroll, G. H. Weiss, I. P. Ipatova, *Theory of lattice dynamics in the harmonic approximation, Solid State Phys.: Advances in research and application* (Academic Press, 2^{nd} Ed, 1971)

[3] A. A. Quong, B. M. Klein, Phys. Rev. B **46**, 10734 (1992).

[4] K. Schmalzl, D. Strauch, H. Schober, Phys. Rev. B **68**, 144301 (2003)

[5] M. T. Dove, *Introduction to Lattice Dynamics* (Cambridge University Press, 1993)

[6] L. Orsingher, G. Baldi, A. Fontana, L. E. Bove, T. Unruh, A. Orecchini, C. Petrillo, N. Violini, F. Sacchetti, Phys. Rev. B **82**, 115201 (2010)

[7] H. G. Bohn, W. Zinn, B. Dorner, A. Kollmar, Phys. Rev B **22**, 5447 (1980)

[8] S. P. Bayrakci, T. Keller, K. Habicht, B. Keimer, Science **312**, 1926 (2006)

[9] G. G. Low, A. Okazaki, R. W. H. Stevenson, K. C. Turberfield, J. appl. Phys. **35**, 998 (1964)

B 5 Crystal Field Effects and Excitations[1]

Marjana Ležaić and Martin Schlipf
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Crystal field theory (CFT), although quite simple, has proven to be a powerful tool in description of *coordination compounds*, characterized by three-dimensional bonds delocalized around a center (*coordination bonds*). The delocalization of the bonds is realized via the d or f orbitals of the central atom, which have many lobes differently oriented in space. In contrast to these, the s and p orbitals can provide only localized or linearly delocalized bonds [1].
CFT can be considered a decedent of the electrostatic theory, developed at the beginning of 20th century by Kossel [2] and Magnus [3]. The theory assumes that the central atom and the ligands are bound together by the ion-ion or ion-dipole electrostatic interactions. The same assumption is carried over to the CFT but in the latter, additionally, the quantum-mechanical nature of the system is taken into account. In 1929, Bethe grounded the CFT in his description of the term splitting in crystals [4]. The theory was further developed in 1930s by Van Vleck and his collaborators [5], who studied the magnetic properties of transition-metal ions in crystals. In 1950s, CFT was further developed and used in order to explain the origin of colours of coordination compounds as well as a number of their other optical, thermodynamic, magnetic and electric properties (see, e.g. Refs. [6, 7]).

Basic statements of the crystal field theory

CFT relies on three main assumptions:

- Stability of a coordination compound is a result of the electrostatic interaction between the central atom or ion and ligands, which can be ions or dipoles.
- Ligands are considered to be structureless sources of electrostatic fields, but a detailed electronic structure of the central atom is taken into consideration.
- The whole system is described by means of quantum mechanics.

Following the assumptions above, the electronic structure of the system is described by the Schrödinger equation with Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + V + W, \tag{1}$$

where $\mathcal{H}_0$ includes the kinetic energy of n electrons of the central atom and their interaction with its nucleus. V stands for interaction of the electrons of the central atom with N ligands taken as point charges q_i, at positions $\mathbf{R}_i(R_i, \theta_i, \varphi_i), i = 1, ...N$ (the nucleus of the central atom is taken to be at the origin),

$$V = -\sum_i^N \sum_j^n \frac{|e|q_i}{\mathbf{r}_j - \mathbf{R}_i}, \tag{2}$$

and W is the the electrostatic interaction of the ligands and the nuclear charge of the central atom (amounting to $Z|e|$, where Z is the atomic number and e is the charge of one electron),

$$W = \sum_i^N \frac{Z|e|q_i}{R_i}. \tag{3}$$

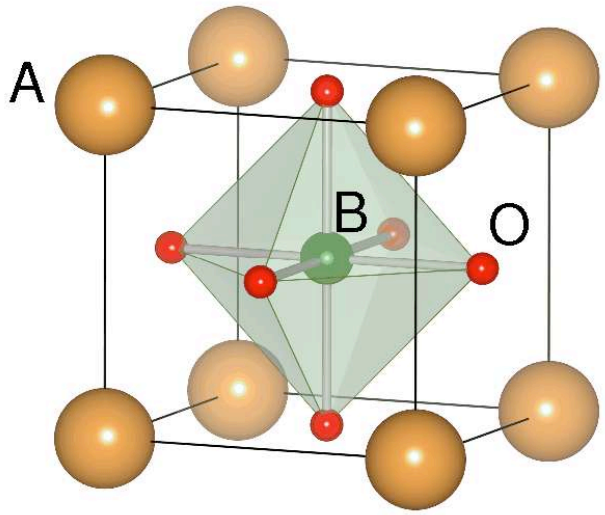


Fig. 1: *Perovskite structure, ABO_3; A and B-sites are occupied by cations. B-site cation (which can be, for example, a transition-metal element) is octahedrally coordinated by six oxygen anions [8].*

For ligands carrying negative charge ($q_i < 0$), the term (3) is negative, providing the stabilization energy for the complex. In this case, however, V is positive, i.e. the electron-ligand interaction destabilizes the complex. This is the case for many real systems. For instance, in transition-metal oxides in perovskite structure, ABO_3 (Fig. 1), the central atom is the transition-metal cation at the B site, octahedrally coordinated by 6 oxygen anions. One example of a perovskite is $LaMnO_3$: here, Mn^{3+} cation is at the B-site and the O^{2-} anions are ligands that are, in spirit of the CFT, treated as structureless negative point charges. Together they form an octahedral complex. The ligands can also be more complex, like in another octahedral complex, hexaaquotitanium (III), $[Ti(H_2O)_6]^{3+}$, but in CFT they will be treated as point charges. The common point of the two octahedral complexes is that the ligands are bound to the central, metal ion by electron pairs contributed by each ligand, located at the corners of an octahedron centered around the metal ion: these electron pairs are equivalent to clouds of negative charge that are directed from near the metal ion out toward the corners of the octahedron. This is an octahedral *ligand field* (also called *crystal field*).

In the following, we will study the main consequence of the ligand field in coordination complexes: a splitting of the electronic energy levels of the central atom, that are, when the atom is isolated, degenerate. The basic ideas will be demonstrated on octahedral complexes, but similar considerations can be straight-forwardly applied to other geometries. For a simpler demonstration of the crystal-field effects, we will consider mainly the transition-metal complexes, where the spin-orbit coupling on the central atom can be considered weak compared to the crystal field. The treatment for a general case of complexes where the spin-orbit interaction is non-negligible will be outlined in Section 5.

Note that although surprisingly powerful, the CFT is limited by its assumptions: the properties of coordination compounds that can be analyzed with the use of CFT are the ones originating in the modifications of the electronic structure of the central atom in the field of ligands. Properties as e.g. complex formation, reactivity, ligand activation etc., that depend on the nature of bonds and charge distribution, require more sophisticated (and necessarily more complicated) theories.

2 Energy levels of one d electron

Firstly, we will focus on the case of a central atom with only one d electron, excluding in this way additional complications that arise due to the electron-electron interaction in systems with more electrons in partly filled d shell. For this purpose, we will take the previously mentioned hexaaquotitanium (III) as an example. In this complex, Ti ion is in oxidation state +3. Since the outer configuration of Ti is $4s^2 3d^2$, the Ti^{3+} will be a d^1 cation, characterized with an orbital momentum $L = 2$, an orbital degeneracy $2L + 1 = 5$, and a spin $S = 1/2$. The five orbital states have identical radial parts, and five possible angular parts that can be chosen such that they are real functions, linear combinations of spherical harmonics (see Table 1). The lobes (distribution maxima) of the d_{z^2} and $d_{x^2-y^2}$ orbitals are oriented along the coordinate axes (see Fig. 2), while the remaining three (d_{xy}, d_{xz} and d_{yz}) have their lobes oriented towards the space between the coordinate axes.

	Real angular parts of the d-states		
orbital	Polar coordinates	Cartesian coordinates	Y_{lm} combination
d_{z^2}	$\sqrt{5/16\pi}\,(3\cos^2\theta - 1)$	$\sqrt{5/16\pi}\,r^{-2}(3z^2 - r^2)$	$Y_{2,0}$
$d_{x^2-y^2}$	$\sqrt{15/16\pi}\,\sin^2\theta\cos(2\varphi)$	$\sqrt{15/16\pi}\,r^{-2}(x^2 - y^2)$	$Y_{2,2} + Y_{2,-2}$
d_{xy}	$\sqrt{15/16\pi}\,\sin^2\theta\sin(2\varphi)$	$\sqrt{15/16\pi}\,r^{-2}xy$	$Y_{2,2} - Y_{2,-2}$
d_{xz}	$\sqrt{15/4\pi}\,\sin\theta\cos\theta\cos\varphi$	$\sqrt{15/4\pi}\,r^{-2}xz$	$Y_{2,1} + Y_{2,-1}$
d_{yz}	$\sqrt{15/4\pi}\,\sin\theta\cos\theta\sin\varphi$	$\sqrt{15/4\pi}\,r^{-2}yz$	$Y_{2,1} - Y_{2,-1}$

Table 1: *Orthonormalized real angular parts of one-electron d-functions, $Y_{lm}(\theta, \varphi)$*

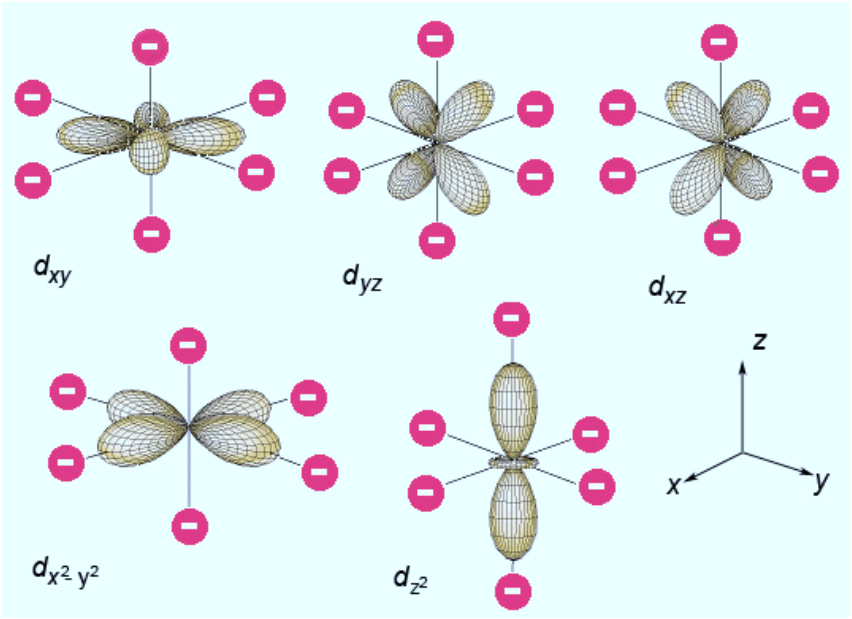


Fig. 2: *d-orbitals in octahedral ligand field. Upper row: t_{2g} orbitals, lower row: e_g orbitals. Pink circles show the positions of negative ligands.*

Following the rules of the group theory, we can already anticipate the effect an octahedral ligand field will have on the d-energy levels of the central ion. A free ion possesses full rotational symmetry. The irreducible representations of this symmetry group are generated by the spherical harmonics $Y_{lm}(\theta, \varphi)$ and are $(2l+1)$-dimensional. For $l = 2$, we have a five-dimensional irreducible representation, corresponding to $m = -2, -1, 0, 1, 2$. Placing the ion in a crystal environment means that we are now dealing with a lower symmetry, namely the crystal point group. However, the irreducible representations of a symmetry group are generally *reducible* representations of its subgroup. We therefore expect that upon reducing the symmetry, the $(2l+1)$-fold degeneracy of each level will in general be partially lifted. A detailed group-theoretical analysis can be found in Ref. [9]. Here we will just state its final result: In an octahedral crystal field (point-group symmetry O_h), the atomic 5-fold degenerate D term splits into a 2-fold degenerate E_g and a 3-fold degenerate T_{2g} term (Mulliken notation is used [10]). Two (degenerate) electronic states with the wavefunctions transforming according to the irreducible representation E_g are usually denoted as e_g. These are the d_{z^2} and $d_{x^2-y^2}$ states. The wavefunctions of the remaining, threefold degenerate, t_{2g} states (d_{xy}, d_{xz} and d_{yz}) transform according to the T_{2g} irreducible representation.

The outlined group-theoretical approach can be understood from the simple considerations involving electrostatic interactions of the electronic cloud with the ligand field. Let us examine what happens with the energy of the d-levels of the metal ion in an octahedral surrounding of negatively charged ligands. This situation is depicted in Fig. 2. Due to the different shape of the electron distribution of the five d-orbitals, we expect that the ligands surrounding the central ion will influence the electron differently depending on the type of the orbital it occupies. It is easy to see that the electron experiences stronger electrostatic repulsion by the negative ligands when

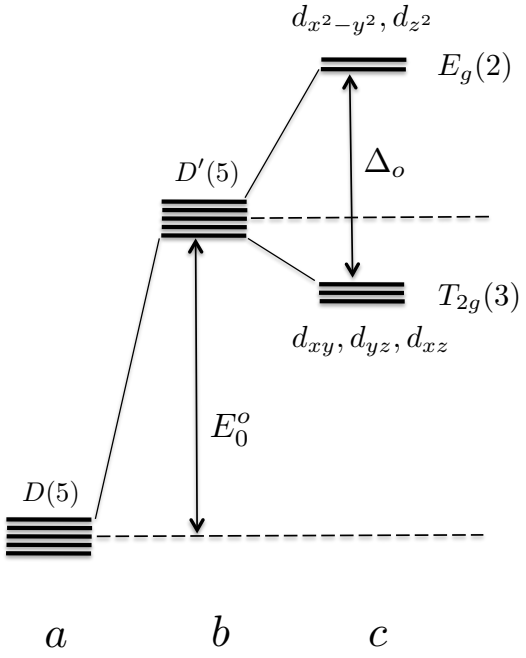


Fig. 3: *Splitting of the d-states in octahedral ligand field: (a) atomic D term; (b) destabilization by the ligand charges when they are assumed to have a symmetric distribution around the metal ion; (c) splitting into a doubly degenerate term with E_g symmetry and a triply degenerate term with T_{2g} symmetry.*

it is in one of the e_g orbitals, oriented directly towards the ligands. This means that the energy of the electron in these orbitals becomes higher with respect to the t_{2g} ones, i.e. the d-levels are now *crystal-field split*.
Notice that not only the energy of e_g levels rises with respect to that of the isolated ion; in fact all five d-states are destabilized in the ligand field, but the amount of the destabilization depends on the spatial orientation of the orbitals. This destabilization is compensated by the electrostatic attraction between the core of the metal ion and the ligands (Eq. (3)). The whole effect can be imagined as consisting of two steps: destabilization and splitting (see Fig. 3). The destabilization energy E_0 corresponds to the repulsion of the electrons of the metal ion by the ligand charges when they are assumed to have a symmetric distribution around the metal ion. Splitting Δ is, for a given central atom and ligands, geometry-specific (in Fig. 3 we have $\Delta = \Delta_o$ for octahedral ligand field).

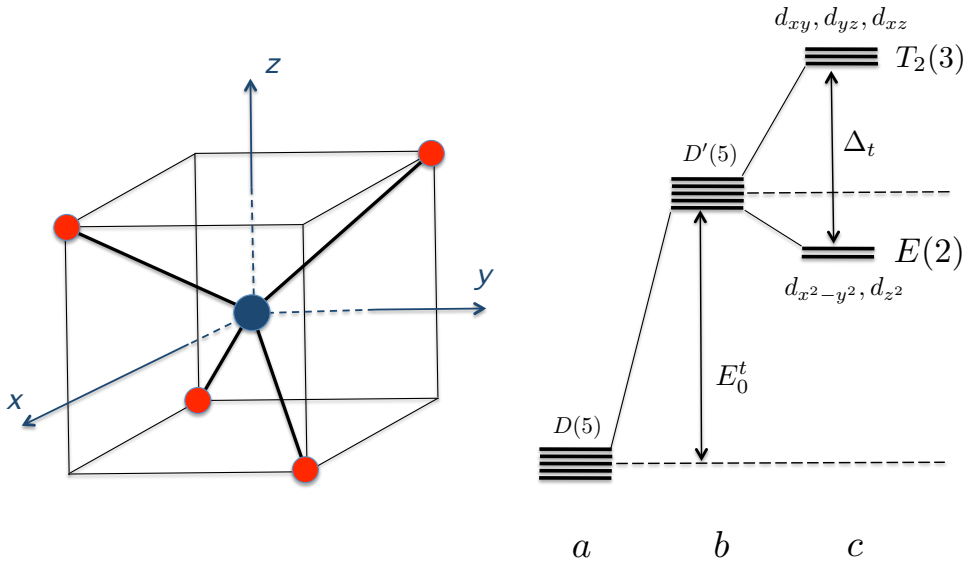


Fig. 4: *Left: An atom in a tetrahedral ligand environment. Right: Splitting of the d-states in tetrahedral ligand field: (a) atomic D term; (b) destabilization by the ligand charges when they are assumed to have a symmetric distribution around the metal ion; (c) splitting into a doubly degenerate term with E symmetry and a triply degenerate term with T_2 symmetry. Note that the sequence of the energies of the split levels is reversed with respect to that in an octahedral ligand field (Fig. 3).*

Environments of different symmetry will cause not only splittings of different size, but also the split levels will in general be ordered differently. For example, in a tetrahedrally coordinated central atom, the d-orbitals will also split into a set of 3-fold degenerate t_2 and 2-fold degenerate e ones, but in this case the e orbitals will be the ones with lower energy. To understand this, imagine four ligands lying at alternating corners of a cube to form a tetrahedral geometry (Fig. 4 left). The $d_{x^2-y^2}$ and d_{z^2} orbitals on the metal ion at the center of the cube lie between the ligands, and the d_{xy}, d_{xz}, and d_{yz} orbitals point toward the ligands. As a result, the splitting observed in a tetrahedral crystal field is the opposite of the splitting in an octahedral complex (Fig. 4 right). Note that a tetrahedral complex has fewer ligands than the octahedral one causing the destabilization energy E_0^t and the magnitude of the splitting Δ_t to be smaller (given the same ligand charges and their distances to the central atom): $E_0^t = \frac{2}{3}E_0^o$, $\Delta_t = \frac{4}{9}\Delta_o$. Splitting of the d-levels in further crystal fields of different symmetries is illustrated in Fig. 5.

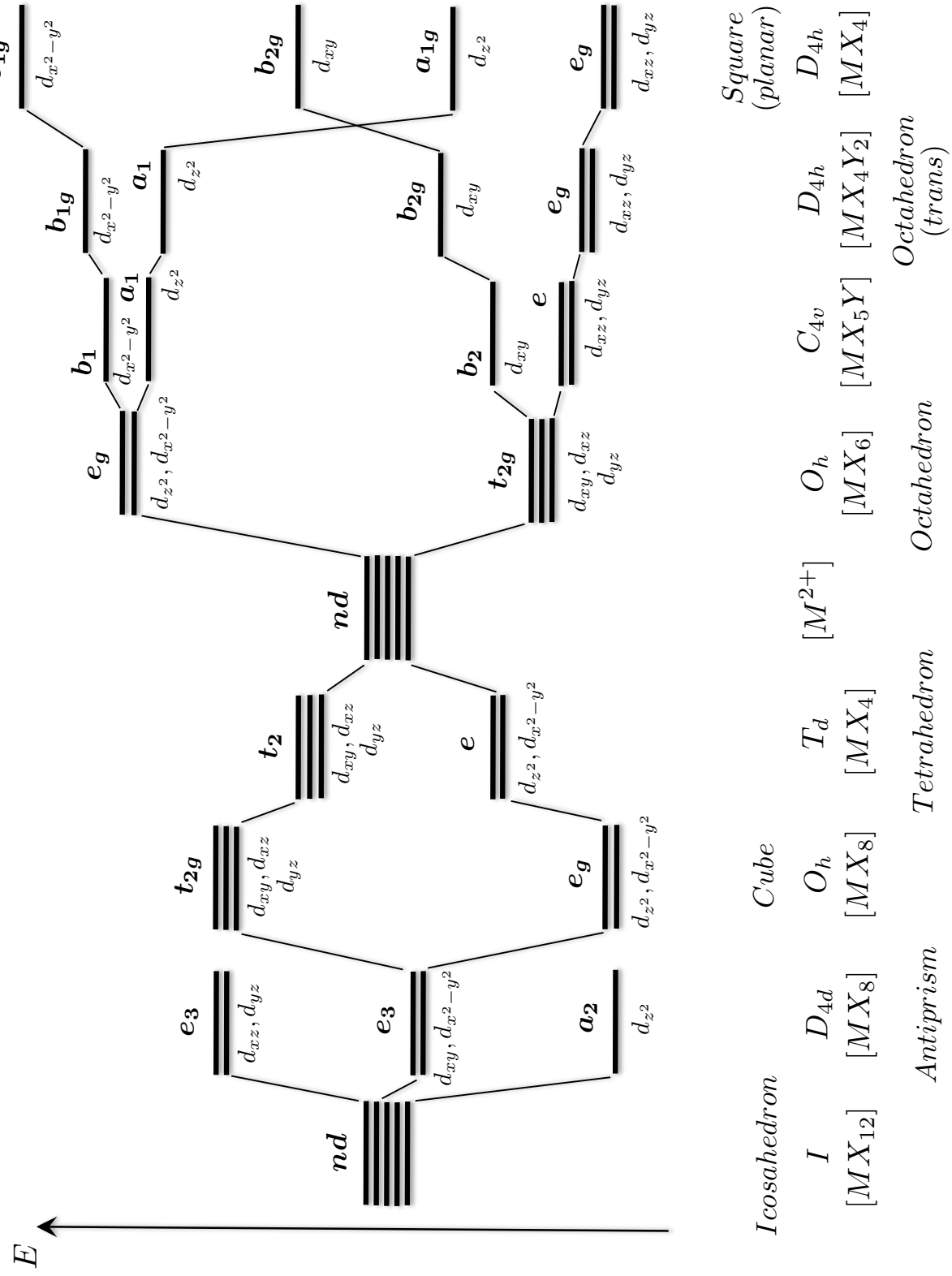


Fig. 5: *Splitting of the* d*-states in ligand fields of different symmetries. Adopted from Ref. [1].*

Let us mention here an effect that we will come back to in Section 7. We have seen that the d-levels, 5-fold degenerate in an isolated atom, split into levels with smaller degeneracy under the influence of a ligand field. Additional lifting of degeneracy of these levels is achieved if the structure of the complex distorts. For example, consider the case of a tetragonally distorted octahedron, elongated along the z-axis. In this case, the energies of the $d_{x^2-y^2}$ and d_{z^2} orbitals are no longer equal, because now the electron experiences smaller electrostatic repulsion from the ligands if it is in the d_{z^2} orbital. In the same way, the repulsion is weaker if the electron is in d_{xz} or d_{yz} orbital, compared to the d_{xy} one. Therefore, in a tetrahedrally distorted octahedral ligand field, the atomic 5-fold degenerate d-levels split into four terms of which only the d_{xz} and d_{yz} remain degenerate. The remaining degeneracy can further be lifted if the ligand-field symmetry is lowered, such that the two directions, x and y, are no more equivalent.
In a given geometry, the crystal-field splitting Δ will depend both on the central atom and on the ligands. The range of values for a given geometry is large. For example, its value is 100 kJ/mol in the $[Ni(H_2O)_6]^{2+}$ ion, and 520 kJ/mol in the $[Rh(CN)_6]^{3-}$ ion. As a consequence, the transition-metal complexes display a variety of colours. For example, from previous considerations one can easily explain the origin of the violet colour of hexaaquotitanium (III) complex. Here, the only d-electron of the Ti^{3+} ion occupies, in the ground state, one of the lower-lying t_{2g} states. The $t_{2g} - e_g$ splitting amounts to $\Delta = 240$ kJ per mole, which corresponds to the light of blue-green colour. Absorption of this light puts the complex into an excited state by promoting the electron to the higher e_g orbitals. When a solution of $[Ti(H_2O)_6]^{3+}$ is illuminated with white light, the blue-green light is in this way absorbed, and the solution appears violet.
From the studies of absorption spectra of transition-metal complexes, the following observations characterizing the size of d-level splitting, known as the *spectrochemical series*, were made:

- When geometry and ligands are held constant, splitting depends on the metal ion and decreases in the order:

  $$Pt^{4+} > Ir^{3+} > Rh^{3+} > Co^{3+} > Cr^{3+} > Fe^{3+} > Fe^{2+} > Co^{2+} > Ni^{2+} > Mn^{2+}.$$

  Metal ions at the left of the series are called *strong-field ions*, and the ones at the right are known as *weak-field ions*, indicating that larger splitting is caused by stronger crystal fields.

- When geometry and metal ion are held constant, splitting depends on the ligands and decreases in the order:

  $$CO \approx CN^- > NO_2^- > NH_3 > H_2O > OH^- > F^- > SCN^- \approx Cl^- > Br^- > I^-.$$

  Ligands that give rise to a large splitting are called *strong-field ligands*. Those at the opposite extreme are known as *weak-field ligands*.

3 Energy levels of several d electrons

When the central atom contains more than one d electron above a closed shell, the picture described in previous section is complicated by the electron-electron interactions. The consequences of these interactions for the electronic configuration of solids are summarized in Hund's rules. According to these rules, we should firstly fill up all the sub-shell orbitals (e.g. all the degenerate d orbitals) with electrons of the same spin (one electron in each orbital). In this way,

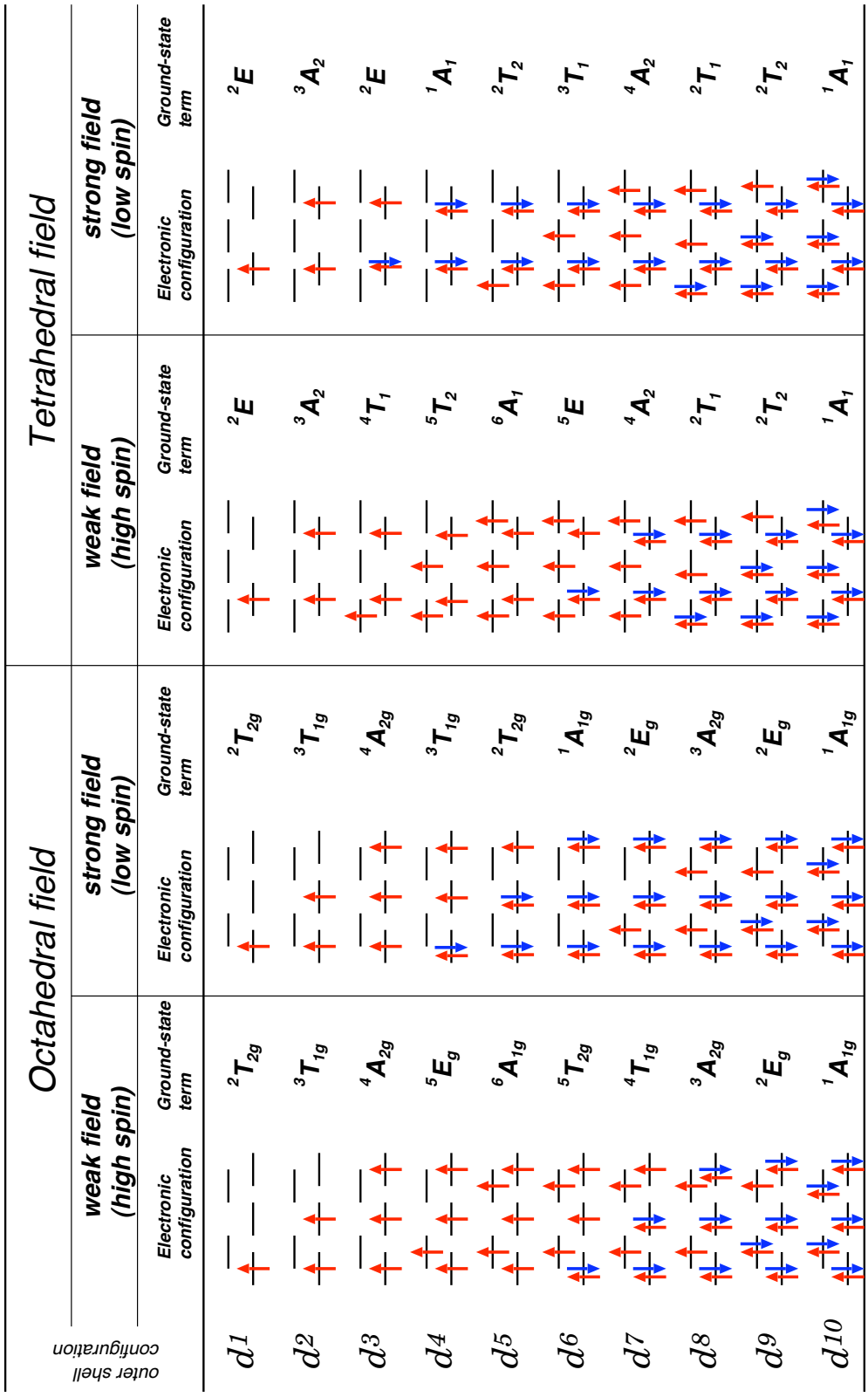


Table 2: *Electronic d-configurations in weak and strong octahedral and tetrahedral ligand fields. Three (two) lines at the same level depict the threefold (twofold) degenerate t_{2g} (e_g) orbitals.*

the electrons that besides the Coulomb repulsion also feel the repulsion due to the same spin-state are furthest apart. Only then should we start pairing the electrons of the opposite spins in the same orbitals.

As we have seen, in a crystal field the electrons prefer the orbitals in which they feel a weaker repulsion from the ligands (e.g. the t_{2g} orbitals in an octahedral crystal field). When the crystal-field splitting is weak, Hund's rules are still valid and the orbitals are filled up to the highest spin state. These are the so-called *high-spin complexes*. However, for sufficiently large splitting, it becomes energetically cheaper to firstly fully occupy the set of crystal-field split orbitals with lower energy and then continue with the filling of the higher ones. For example, consider a central atom with 7 d-electrons in a strong octahedral crystal field. If the $t_{2g} - e_g$ splitting Δ_o is larger than the energy needed to pair two electrons in the same orbital, this will result in an electronic configuration where three pairs of electrons occupy the t_{2g} orbitals and one electron is placed in one of the higher e_g orbitals. In this case we are dealing with a *low-spin complex*. In Table 2, the electronic d configurations in octahedral and tetrahedral weak and strong crystal fields are tabulated (along with the ground-state symmetry).

A metal ion in a given oxidation state can form complexes with different ligands; depending on the position of the ligands in spectrochemical series, the same metal ion can appear either in a low-spin or in a high-spin state in different complexes. An interesting case occurs when the ligand field is of intermediate strength. An example is a perovskite, $LaCoO_3$. In this material the difference in the high-spin and low-spin configurations is very small (0.08 eV, according to Ref. [11]), the low-spin one being the ground state. In this case, the low-spin configuration corresponds to $t_{2g}^6 e_g^0$ and the high-spin one to $t_{2g}^4 e_g^2$. With an increase in temperature, the high-spin state starts being populated. This is reflected in the magnetic susceptibility, which shows a maximum around 90 K and a Curie-Weiss-like decrease at higher temperatures [12]. It was also suggested [13] that in fact there is an additional, intermediate spin-state with a configuration $t_{2g}^5 e_g^1$, which is also orbitally ordered.

4 Magnetic moment in a crystal field

Following Hund's rules, the magnetic moment of a free atom μ_J can be easily estimated from

$$\mu_J = g_J \mu_B \sqrt{J(J+1)}, \tag{4}$$

where μ_B is a Bohr magneton, J is the total angular momentum $J = L + S$ (L-orbital angular momentum and S-spin), and

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \tag{5}$$

is the Landé g-factor. For a free electron, $L = 0$, $J = S$ and $g_J = 2$. In a special case of a relatively weak spin-orbit coupling (as in transition metals of the first row), the effective measured magnetic moment can be expressed as

$$\mu_{eff} = \mu_B \sqrt{4S(S+1) + L(L+1)}. \tag{6}$$

However, the experimental data for transition-metal ions in complexes show that, in most cases, the magnetic moment is closer to the value

$$\mu_{eff} = 2\mu_B \sqrt{S(S+1)}, \tag{7}$$

i.e, as if the orbital moment L vanished. This effect is called *quenching of the orbital moment* by the crystal field. Its origin will become clear from the following simple example [14].

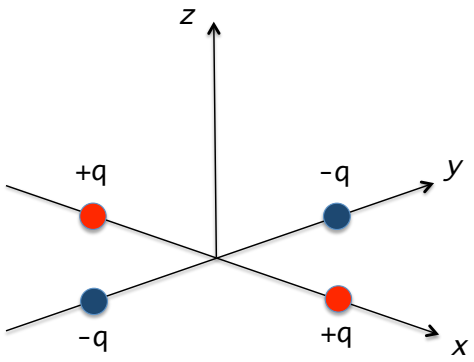


Fig. 6: *A hypothetical ligand field. We are investigating its influence on the central atom, positioned at the origin.*

Consider an atom with one p electron above a closed shell. For simplicity, we will neglect the electron's spin. To describe the three degenerate p functions we can chose either the set of functions

$$|x\rangle = xf(r), \qquad |y\rangle = yf(r), \qquad |z\rangle = zf(r), \tag{8}$$

or

$$|+\rangle = \frac{1}{\sqrt{2}}(x+iy)f(r), \qquad |z\rangle = zf(r), \qquad |-\rangle = \frac{1}{\sqrt{2}}(x-iy)f(r). \tag{9}$$

The latter are the eigenstates of the z-component of angular momentum L_z, with eigenvalues $m = +1, 0, -1$ (from left to right).

Let us now assume that the atom is placed in a crystal field depicted in Fig. 6, i.e. in the center of a square formed by four charges lying at equivalent distances from the atom, two positive ones lying on the x-axes and two negative ones lying on the y-axis. It is easy to convince oneself that in a perturbing potential V_p with such a symmetry, the eigenstates are described only with the first set of functions, where $\langle x|V_p|x\rangle = \langle y|V_p|y\rangle = \langle z|V_p|z\rangle = 0$. Orbital degeneracy is lifted: since the p_x orbital extends towards positive charges its energy is lowered, while the energy of p_y orbital, extending towards negative charges, is raised; the energy of the p_z orbital is in between. The quenching of the orbital moment can, somewhat hand-wavingly, be understood as follows. The ground state $|x\rangle$ can be written as a a linear combination of states $|+\rangle$ and $|-\rangle$,

$$|x\rangle = \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(x+iy)f(r) + \frac{1}{\sqrt{2}}(x-iy)f(r)\right] = \frac{1}{\sqrt{2}}\left[|+\rangle + |-\rangle\right]. \tag{10}$$

The states $|+\rangle$ and $|-\rangle$, with $m = +1$ and $m = -1$ respectively, are states with opposite circulation of electron about the z-axis, resulting in a zero total circulation in a linear combination with equal coefficients. This, in other words, implies a zero expectation value of the angular momentum component L_z in the ground state $|x\rangle$.

In a more formal manner, we can demonstrate that the same holds for any *real* wavefunction $|\Psi\rangle$ by directly calculating the expectation value of the components of angular momentum operator. For example, consider the L_z component,

$$L_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{11}$$

It is clear from this expression that the expectation value $\langle\Psi|L_z|\Psi\rangle$ must be purely imaginary. However, L_z is Hermitian, and the diagonal matrix elements of any Hermitian operator must be real. Thus, $\langle\Psi|L_z|\Psi\rangle$ must be real. The only possibility satisfying both conditions is that $\langle\Psi|L_z|\Psi\rangle$ vanishes. The same can easily be shown for all three components of orbital angular momentum operator:

$$\langle\Psi|L_x|\Psi\rangle = \langle\Psi|L_y|\Psi\rangle = \langle\Psi|L_z|\Psi\rangle = 0. \tag{12}$$

We therefore conclude that when the wavefunction describing a state of the system is real, the angular momentum in that state is quenched.

Now we should answer the arising question: when can a state of the system be described by a real function? Let Ψ be a *non-degenerate* eigenstate of the system. If spin-dependent interactions are neglected, the Hamiltonian H is a real operator. In this case, the eigenfunction Ψ is also real: if it was complex, $\Psi = \psi_1 + i\psi_2$ and H real, than both ψ_1 and ψ_2 would also have to be eigenfunctions of the Hamiltonian with the same eigenvalue, i.e. the level would have to be degenerate, in contradiction with the initial assumption. We came to the conclusion that, if in the crystal field a state becomes non-degenerate, the orbital angular momentum of that state is quenched.

What happens with the orbital moment when the eigenfunction $|\Psi\rangle$ is degenerate? This question can be answered through an analysis by means of group theory. Without going into details (which can be found in e.g. Ref. [1]), we will just state that for a degenerate electronic state $|\Psi\rangle$ of a cubic system, orbital magnetism is quenched in all cases except when $|\Psi\rangle$ transforms according to either T_1 or T_2 irreducible representation of the octahedral point group O_h. This conclusion, according to Table 2, means that in octahedral complexes an orbital contribution to the measured magnetic moment μ_{eff} can be expected for the electronic configurations $d^1,\ d^2$, low-spin $d^4,\ d^5$ and high-spin $d^6,\ d^7$.

Table 3 shows how the measured magnetic moments of several ions in complexes, compare to that expected from the Eqns. (6) and (7), reflecting the crystal-field induced quenching of orbital moments.

ion	shell	S	L	J	μ_1	μ_{exp}	μ_2
Ti^{3+}, V^{4+}	$3d^1$	$\frac{1}{2}$	2	$\frac{3}{2}$	1.55	1.70	1.73
V^{3+}	$3d^2$	1	3	2	1.63	2.61	2.83
Cr^{3+}, V^{2+}	$3d^3$	$\frac{3}{2}$	3	$\frac{3}{2}$	0.77	3.85	3.87
Mn^{3+},Cr^{2+}	$3d^4$	2	2	0	0	4.82	4.90
Fe^{3+}, Mn^{2+}	$3d^5$	$\frac{5}{2}$	0	$\frac{5}{2}$	5.92	5.82	5.92
Fe^{2+}	$3d^6$	2	2	4	6.70	5.36	4.90
Co^{2+}	$3d^7$	$\frac{3}{2}$	3	$\frac{9}{2}$	6.63	4.90	3.87
Ni^{2+}	$3d^8$	1	3	4	5.59	3.12	2.83
Cu^{2+}	$3d^9$	$\frac{1}{2}$	2	$\frac{5}{2}$	3.55	1.83	1.73
Zn^{2+}	$3d^{10}$	0	0	0	0	0	0

Table 3: *Magnetic moments μ_1 of $3d$ ions in μ_B, calculated from Hund's rules Eq. (4), compared to the measured experimental values μ_{exp} and the moment μ_2 obtained from Eq. (7), where the orbital quenching is assumed. Adopted from Ref. [15]*

5 Crystal field and spin-orbit coupling

So far we have considered only electrostatic interactions in complexes. We assumed that there was no external magnetic field and that the spin-orbit interaction was negligible. Let us now examine possible additional contributions we left out in the initial Hamiltonian, Eq. (1), and write the full Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + V_{CF} + V_M, \tag{13}$$

where V_{CF} is the crystal field contribution ($V_{CF} = V + W$, Eqns. (2) and (3)) and V_M is the magnetic contribution, consisting of the spin-orbit interaction and the Zeeman energy:

$$V_M = \lambda \mathbf{L} \cdot \mathbf{S} + \mu_B \mathbf{H} \cdot (2\mathbf{S} + \mathbf{L}), \tag{14}$$

λ being the spin-orbit coupling strength and $\mathbf{H}$ an external magnetic field. It is instructive to compare the energies of the terms in Hamiltonian (13). While the potential and kinetic energy term contained in H_0 amount to 8000-80000 cm^{-1} (1-10 eV), the crystal-field term V_{CF} contributes by 100-10000 cm^{-1} (12.5 meV-1.25 eV). The energy of the spin-orbit interaction ranges between 10 cm^{-1} and 2000 cm^{-1} (1.25-250 meV), while the Zeeman interaction contribution is smaller than 10 cm^{-1} (1.25 meV).
The Zeeman interactions are almost always much smaller than the crystal-field and spin-orbit interactions. The latter two, however, can have comparable energies. This means that, in principle, the V_{CF} and $\lambda \mathbf{L} \cdot \mathbf{S}$ have to be diagonalised together. However, we can separate two extreme cases:

- In heavy rare earth elements, the spin-orbit coupling is strong. In addition to this, the f electrons are usually screened from the crystal field by the outer $s,\ p,\ d$ electrons and are less affected by it than are the d electrons. This is, therefore, the case when the spin-orbit coupling is much stronger than the crystal field. Hence, here we can start from the atomic terms characterized with the quantum number $J = L + S$ and treat the crystal field as a perturbation of these states.

- In contrast to the former case, in this chapter we concentrated mostly on the case of weak (compared to the crystal field) spin-orbit interaction which is true for the light, $3d$-row transition metals. In this case, the crystal field problem is solved first and the spin-orbit interaction is treated as a perturbation.

We will continue with the latter case and include the magnetic contribution (14) in a perturbative manner, assuming the ground state of the unperturbed hamiltonian (1) to be non-degenerate. This might seem as a somewhat arbitrary choice, since the crystal field, as we have seen so far, need not lift the degeneracy of the atomic term completely. However, in the next section we will see that the remaining degeneracy is lifted by distortions of the lattice (this is known as the *Jahn-Teller effect*). From the results of Sec. 4, we conclude that the orbital moment in the ground state is quenched. Spin is completely independent of the orbital part, with a degeneracy $2S + 1$. The eigenfunctions corresponding to the eigenvalues E_n of the unperturbed Hamiltonian can be represented as products of spin states and orbital wave functions. The latter will be labeled $|n\rangle$. The ground-state eigenvalue and the corresponding orbital wave function will be labeled E_0 and $|0\rangle$, respectively. To second order, we get an effective Hamiltonian where only the operator of spin remains:

$$\mathcal{H}_{eff} = \sum_{i,j} \left[2\mu_B H_i(\delta_{i,j} - \lambda\Lambda_{i,j})S_j - \lambda^2 \Lambda_{i,j} S_i S_j - \mu_B^2 \Lambda_{i,j} H_i H_j \right], \tag{15}$$

where $i,\ j$ run over the three components $x,\ y,\ z$ and

$$\Lambda_{i,j} = \sum_n \frac{\langle 0|L_i|n\rangle\langle n|L_j|0\rangle}{E_n - E_0}. \tag{16}$$

The first term of the Hamiltonian (15) represents an effective Zeeman energy, with the scalar Landé factor g substituted by the tensor

$$g_{i,j} = 2(\delta_{i,j} - \lambda\Lambda_{i,j}), \tag{17}$$

where the additional factor $-2\lambda\Lambda_{i,j}$ is the induced orbital moment arising from the mixing with the excited states due to the LS coupling. Owing to the directional dependence of the g-factor, its measurement can give valuable information about the crystal structure.
The second term in the effective Hamiltonian is the anisotropy term, $\mathcal{H}_a$. It is responsible for lifting of the spin-degeneracy, i.e. it introduces an anisotropy energy for different spin directions. The tensor Λ is a real, symmetric tensor. It therefore has three orthogonal eigenvectors (the principal axes). Choosing x, y and z along the principal axes, we can express the tensor in a diagonal form, with the diagonal components Λ_{xx}, Λ_{yy} and Λ_{zz}. The anisotropy term then becomes

$$\begin{aligned} \mathcal{H}_a = -\lambda^2 \Bigg\{ & \frac{1}{3}(\Lambda_{xx} + \Lambda_{yy} + \Lambda_{zz})S(S+1) \\ & + \frac{1}{3}\left[\Lambda_{zz} - \frac{1}{2}(\Lambda_{xx} + \Lambda_{yy})\right]\left[3S_z^2 - S(S+1)\right] \\ & + \frac{1}{2}(\Lambda_{xx} - \Lambda_{yy})(S_x^2 - S_y^2)\Bigg\}. \end{aligned} \tag{18}$$

Let us see how $\mathcal{H}_a$ acts on the $(2S+1)$-fold degenerate states. For this purpose, we need to consider only the terms containing spin-operators. Omitting the constant term we write:

$$\mathcal{H}'_a = DS_z^2 + E(S_x^2 - S_y^2), \tag{19}$$

where

$$D = -\lambda^2\left[\Lambda_{zz} - \frac{1}{2}(\Lambda_{xx} + \Lambda_{yy})\right] \quad \text{and} \quad E = -\frac{\lambda^2}{2}(\Lambda_{xx} - \Lambda_{yy}) \tag{20}$$

For a state characterized with the quantum number M_s, $S_z|M_s\rangle = M_s|M_s\rangle$. Therefore, only the diagonal elements $D\langle M_s|S_z^2|M_s\rangle$ of the the first term in Eq. (19) are nonzero. Using the spin raising and lowering operators $S_\pm = S_x \pm iS_y$, the second term can be transformed to read $E(S_+^2 + S_-^2)/2$, where it becomes clear that only its off-diagonal elements $E\langle M'_s|S_+^2 + S_-^2|M_s\rangle/2$ between the states with $M'_s - M_s = \pm 2$ are not zero.
It follows from (20) that $D = 0$ and $E = 0$ in a cubic symmetry, i.e. when $\Lambda_{xx} = \Lambda_{yy} = \Lambda_{zz}$. In axial symmetry with z as unique axis, we have $\Lambda_{xx} = \Lambda_{yy}$, i.e. $D = \Lambda_{zz} - \Lambda_{xx}$ and $E = 0$. In this case, the eigenvectors of $\mathcal{H}'_a$, Eq. (19), are just the eigenvectors $|M_s\rangle$ of S_z, with the eigenvalues M_s^2. The initial $(2S+1)$-fold spin degeneracy is partially removed: the states with

$+M_s$ and $-M_s$ are still degenerate. Allowing for a non-zero E, i.e. reducing the symmetry further, below the axial one, lifts the remaining degeneracy if S has an integer value. In case of a half-integer S, the degeneracy of the $\pm M_s$ states remains. This is a consequence of the time reversal symmetry known under the name *Kramers degeneracy*, while the degenerate levels are called *Kramers doublets* [16]. Kramers degeneracy is lifted in a magnetic field by the Zeeman term.
Note that the stability of states with a particular value of M_s depends on the sign of D. If $D > 0$, the most stable level is the one with the smallest value of $|M_s|$, i.e. the preferred spin-orientation is in the x-y plane. When $D < 0$, the most stable state is the one with the maximal $|M_s|$ value (preferred spin-orientation is along the z-axis).
Finally, the third term in (15) comes from the second-order perturbation of the Zeeman energy for the orbital angular momentum. It gives rise to a temperature-independent contribution to the susceptibility (*van Vleck susceptibility*) arising from the admixture of multiplets of higher energy than the ground state.

6 Determining the Landé factor: Electron Spin Resonance

In an external magnetic field, the energy levels of an atom are split (Zeeman effect) and the size of the splitting will depend on the Landé factor (see the first term in the effective Hamiltonian (15) and Eq (17)). When electromagnetic waves within the radio and microwave frequencies are absorbed by the material, transitions between these energy levels occurs. This is the phenomenon that is used in the electron spin resonance (ESR) experiments. In the simple case of an atom in a magnetic field $\mathbf{H}$, the energy levels are determined by the Zeeman splitting

$$\epsilon_m = \mu_B g H m, \tag{21}$$

where $m = -J, ..., J-1, J$ is the projection (on the direction of the magnetic field) of the total momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$, H is the magnetic field magnitude and g is given by the Eq. (5). Obviously, for a given magnetic field, the transitions $\epsilon_{m+1} \rightarrow \epsilon_m$ are completely determined by the Landé factor g:

$$\hbar\omega = \epsilon_{m+1} - \epsilon_m = \mu_B g H, \tag{22}$$

where $\hbar\omega$ is the absorbed quantum of electromagnetic radiation.
For the simple case of an atom, Eq. (5) gives $g = 1$ if $S = 0$ and $L \neq 0$, or $g = 2$ if $L = 0$ and $S \neq 0$. As we have seen in the previous section however, when the atom is in a transition-metal compound, g becomes a tensor (Eq. (17)) that can give us a lot of information about the electronic structure of the compound and the symmetry of the specific crystal field.
An illustrative example is the case of the g-factor modification by the admixture of the ligand states in tetragonally distorted Cu^{2+} complexes with D_{4h} symmetry. [1]. The splitting of the ${}^2D(d^9)$ term of Cu^{2+} for this case is shown in Fig. 7 (remember that a hole in a $d^1 0$ configuration has the level splitting inverse of that of an electron in d^1). The ground-state term B_{1g} corresponds to the atomic orbital $d_{x^2-y^2}$. Hence, its orbital moment is quenched (see Sec. 4). The spin-orbit interaction, however, mixes the ground state with the B_{2g} and E_g which originate in the T_{2g} term that does carry an orbital contribution. Without going into details of the derivation (that can be found in Ref. [1]), we will just state that when the spin-orbit interaction is treated as a perturbation, one can in the second-order perturbation theory obtain the following

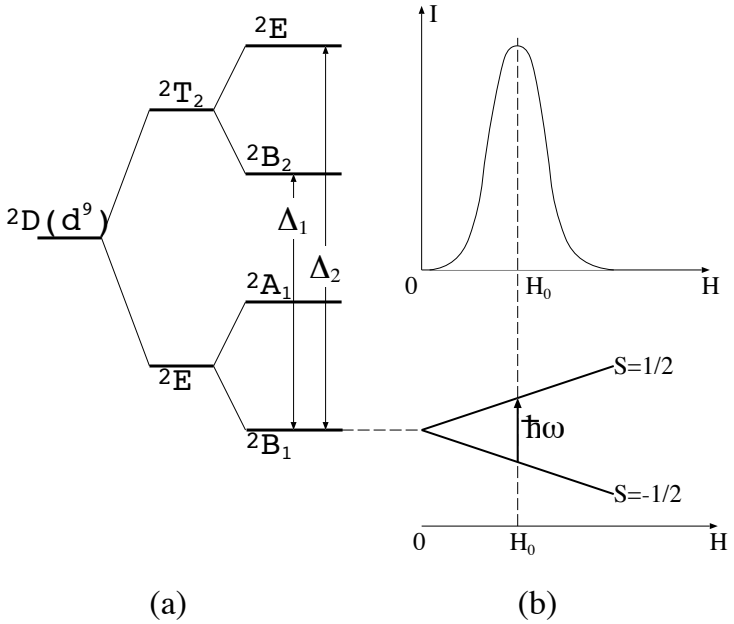


Fig. 7: *(After Ref. [1]) The splitting of the $^2D(d^9)$ term in Cu^{2+} complexes with D_{4h} symmetry (a) and a further splitting in an external magnetic field H (b, lower panel). There is a peak at the resonance absorption at $\mu_B H_0 = \hbar\omega$ (b, upper panel).*

longitudinal (g_l) and normal (g_n) components of the g-tensor (the "longitudinal" and "normal" are defined with respect to the tetragonal axis):

$$g_l = g_{zz} = 2 - \frac{8\lambda}{\Delta_1} N_1^2 N_2^2 \tag{23}$$

$$g_n = g_{xx} = g_{yy} = 2 - \frac{2\lambda}{\Delta_2} N_1^2 N_3^2, \tag{24}$$

where λ is the spin-orbit coupling constant, Δ_1 and Δ_2 are the energy gaps to the excited states depicted in Fig. 7 and N_1, N_2, N_3 are the normalization constants of the wavefunctions of the mixing states B_{1g}, B_{2g} and E_g, respectively, in the so-called weak-covalence model [1]; the normalization constants $N_i = [1 + \gamma_i^2 - 2\gamma_i Si]$ $(i = 1, 2, 3)$ depend on the covalence constant γ_i and the overlap S_i between the states of the the central ion (Cu^{2+}) and those of the ligands.
For pure states of the central ion, $N_1 = N_2 = N_3 = 1$, yielding the orbital contributions of $-8\lambda/\Delta_1$ and $-2\lambda/\Delta_2$ to g_l and g_n, respectively (note that $\lambda < 0$ for Cu i.e. the orbital contributions are positive, increasing the value of the g-factor with respect to that calculated from Eq. (5) for $L = 0$ and $S \neq 0$). If the covalence constants are not zero, the normalization constants are less than 1, reducing the orbital contribution to g_l and g_n $k_l^2 = N_1^2 N_2^2$ and $k_n^2 = N_1^2 N_3^2$ times (the $k_{l,n}^2$ denote the so-called covalence reduction) . This can also be viewed as a reduction of the spin-orbit coupling constant to $\lambda' = k_{l,n}^2 \lambda$, due to the effects of reduced localization by the coordination. λ' can be extracted from the measured g-factors, allowing for an estimation of the covalence reduction.
The described covalence contribution to the g-factor is only one of the possible contributions that influence the ESR spectrum. Further contributions include, e.g. the hyper-fine splitting (due to the interaction with the nuclear spin), superhyperfine splitting (due to the interaction with the nuclear spins of the ligands), paramagnetic relaxation (temperature-dependence of the

line shape due to the interactions with vibrations), vibronic reduction (the reduction of the g-factor due to orbital degeneracy) etc. A detailed analysis of these contributions can be found in Ref. [1].

7 Jahn-Teller effect

As we have seen, a crystal field leads to symmetry-defined lifting of the orbital degeneracy of the atomic term. The resulting ground state need not be non-degenerate. For example, in $LaMnO_3$, the Mn^{3+} ion has a $t_{2g}^3\ e_g^1$ configuration: the one electron in the e_g state can occupy either the d_{z^2} or the $d_{x^2-y^2}$ orbital. However, one should keep in mind that the particular splitting of the d-orbitals into the t_{2g} and e_g ones was derived under the assumption of fixed nuclear positions. "Releasing" the nuclei changes the situation: they can now react to the electronic state, adjusting their position so that a global minimum in energy is reached. We have already described in Sec. 2 the splitting of energy levels of a central atom in a tetragonally distorted octahedron. The distortion is a consequence of the degeneracy of the electronic state. On the example of $LaMnO_3$, this can be understood as follows. Say that the electron placed in e_g states resides in the d_{z^2} orbital. The electron, having negative charge, pushes the negatively charged oxygens lying on the z-axis away, inducing in this way a tetragonal distortion of the system. Due to the distortion, the 2-fold degeneracy of the E term is lifted (now the d_{z^2} orbital has a lower energy than the $d_{x^2-y^2}$ one). Note that, as a byproduct, the energies of the t_{2g} states also change, the one of d_{xy} being higher than that of d_{xz} an d_{yz} which remain degenerate.
The example of $LaMnO_3$ lattice distortion is a consequence of a general phenomenon, described by the *Jahn-Teller theorem*. In its original formulation [17], the theorem states the following:

> The nuclear configuration of any nonlinear polyatomic system in a degenerate electronic state is unstable with respect to nuclear displacements that lower the symmetry and remove the degeneracy.

Jahn-Teller effect is in fact a product of *vibronic interactions* (coupling of electronic and nuclear motion) which take place in systems with a degenerate ground state. Similar changes in nuclear configurations actually do take place also in linear molecules [18] and are referred to as *Renner* or *Renner-Teller* effect. Another effect originating in vibronic interactions is the *pseudo-Jahn-Teller effect*. In this case, however, the vibronic mixing doesn't involve degenerate states, but rather non-degenerate ones that are very close in energy to each other. The pseudo-Jahn-Teller effect is at the microscopic origin of ferroelectricity [19].
In the following, we will see how to find the new nuclear configuration resulting from vibronic interactions. Firstly, let us establish the cases in which the vibronic interaction needs to be taken into account. We consider a system of electrons and nuclei oscillating about their equilibrium positions. Schrödinger equation of the whole system reads

$$\mathcal{H}\Psi(r,Q) = E\Psi(r,Q), \tag{25}$$

where Ψ is the wavefunction of the system and r and Q stand for the whole set of electronic and nuclear coordinates, respectively. We will describe the nuclear positions by means of *symmetrized* displacements, i.e. collective nuclear displacements which under a symmetry operation of the point group of the complex transform according to one of its irreducible representations. Figure 8 shows the symmetrized displacements of an octahedral complex. The exact solution of Eq. (25) is difficult to find. Observing that the mass of nuclei is about 2000 times

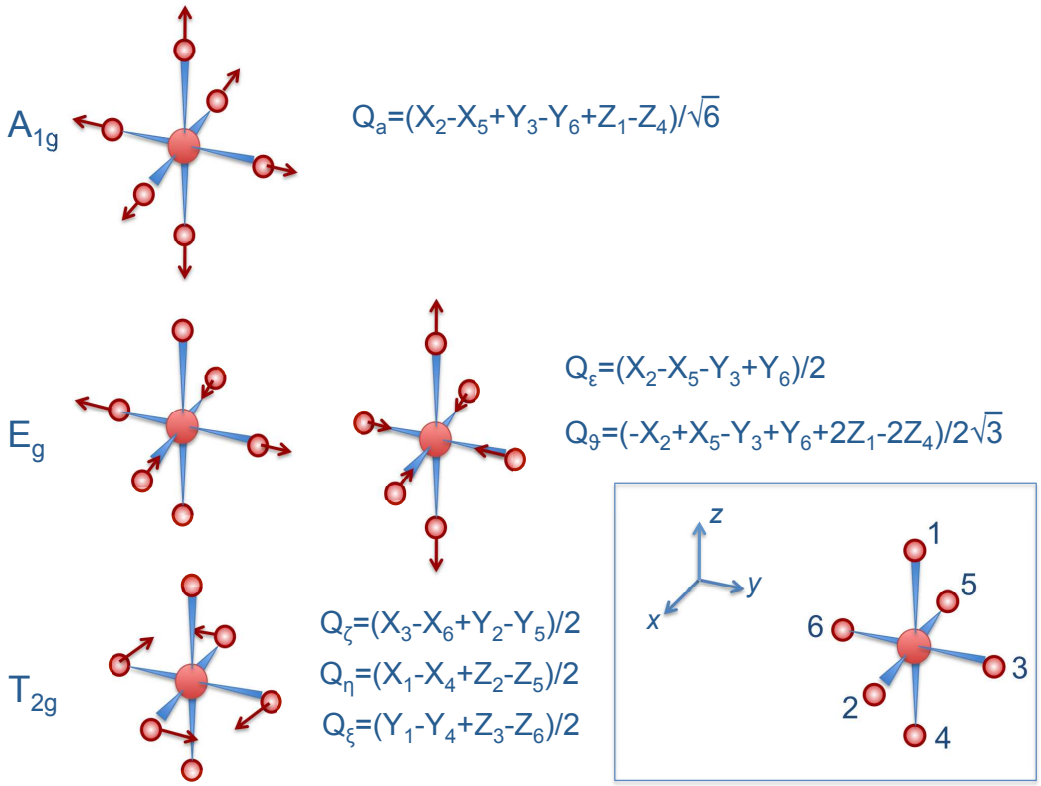


Fig. 8: *Patterns of symmetrized displacements in octahedral complexes labeled by their symmetry (left), and the corresponding expressions in Cartesian coordinates. The inset shows the numbering of atoms used in the expressions and the orientation of coordinate axes.*

larger than that of an electron, we can assume that the electronic motion is much faster than the nuclear one, so that the electronic cloud can be considered as fully relaxed at each instantaneous position of the nuclei. This is the well-known adiabatic or Born-Oppenheimer approximation. Practically, it means that we can firstly solve the electronic part of the Schrödinger equation at each nuclear configuration Q and then use the obtained electronic energy as potential energy for the motion of the nuclei. Let us see when the adiabatic approximation is justified. The full Hamiltonian of the system can be divided into three contributions,

$$\mathcal{H} = \mathcal{H}_r + \mathcal{H}_Q + V(r, Q), \tag{26}$$

where $\mathcal{H}_r$ is the electronic and $\mathcal{H}_Q$ the nuclear part, while $V(r, Q)$ describes the electron-nuclear interaction. Expanding the latter in a series with respect to the symmetrized displacement of nuclei around the high symmetry configuration $Q_\alpha = 0$ we obtain:

$$V(r, Q) = V(r, 0) + \sum_\alpha \left(\frac{\partial V}{\partial Q_\alpha}\right)_0 Q_\alpha + \frac{1}{2}\sum_{\alpha\beta}\left(\frac{\partial^2 V}{\partial Q_\alpha \partial Q_\beta}\right)_0 Q_\alpha Q_\beta + \ldots \tag{27}$$

We now solve the adiabatic part of the Hamiltonian, (26), $\mathcal{H}_\mathcal{A} = \mathcal{H}_r + V(r, 0)$:

$$\mathcal{H}_{\mathcal{A}}\varphi_k(r) = \epsilon_k^{el}\varphi_k(r), \quad k = 1, 2, 3... \tag{28}$$

obtaining a set of electronic wavefunctions $\varphi_k(r)$ and energies ϵ_k^{el}. The total wavefunction $\Psi(r, Q)$ can be expanded in terms of electronic wavefunctions with coefficients $\chi(q)$ depending on the nuclear coordinates:

$$\Psi(r, Q) = \sum_k \chi_k(Q)\varphi_k(r). \tag{29}$$

The coefficients $\chi_k(Q)$ are then found from the system of coupled equations obtained from the substitution of the expansion (29) and Eq. (26) into Eq. (25),

$$\left[\mathcal{H}_Q + \epsilon_k(Q) - E\right]\chi_k(Q) + \sum_{m\neq k} W_{km}(Q)\chi_m(Q) = 0, \quad k, m = 1, 2, 3... \tag{30}$$

Here, $W(r, Q)$ is the *vibronic potential*, consisting of the terms in expansion (27) that depend on the nuclear coordinates,

$$W(r, Q) = \sum_\alpha \left(\frac{\partial V}{\partial Q_\alpha}\right)_0 Q_\alpha + \frac{1}{2}\sum_{\alpha\beta}\left(\frac{\partial^2 V}{\partial Q_\alpha \partial Q_\beta}\right)_0 Q_\alpha Q_\beta + ... \tag{31}$$

and

$$\epsilon_k(Q) = \epsilon_k^{el} + W_{kk}(Q) \tag{32}$$

is the potential energy of the nuclei in the mean field of the electrons in the state $\varphi_k(r)$. If the state $\varphi_k(r)$ is non-degenerate and there are no other electronic states with energy very close to $\epsilon_k^{el}(Q)$, then $\epsilon_k(Q)$ is called *the adiabatic potential energy surface* (APES).
The only term in Eq. (30) that, through a nuclear motion, mixes different electronic states is the sum over the non-diagonal elements of the vibronic interaction potential, $\sum_{k\neq m} W_{km}(Q)$. If we assume that this mixing can be neglected ($W_{km}(Q) = 0$ for $k \neq m$), we obtain from (30) a system of decoupled equations

$$\left[\mathcal{H}_Q + \epsilon_k(Q) - E\right]\chi_k(Q) = 0, \quad k = 1, 2, 3... \tag{33}$$

describing the nuclear motion in the mean field $\epsilon_k(Q)$ of electrons in the states $\varphi_k, k = 1, 2, 3....$
For each (non-degenerate) k in the harmonic approximation, using normal coordinates Q_α for the nuclear displacements, the APES can be written in the form (we omit the subscript k for simplicity)

$$\epsilon(Q) = \frac{1}{2}\sum_\alpha K_\alpha Q_\alpha^2, \tag{34}$$

while the kinetic energy term has the form

$$\mathcal{H}_Q = -\frac{\hbar^2}{2}\sum_\alpha \frac{1}{M_\alpha}\frac{\partial^2}{\partial Q_\alpha^2}, \tag{35}$$

where K_α is the force constant and M_α is the reduced mass of the normal vibration α. Eqns. (33) then transform into the well-known harmonic oscillator equations, yielding the eigenvalues

$$E_{n_\alpha} = \hbar\omega_\alpha(n_\alpha + \frac{1}{2}), \qquad n_\alpha = 0, 1, 2... \tag{36}$$

where ω_α is the frequency of the normal vibration α. It becomes intuitively clear that the question whether or not the vibronic interactions that mix different electronic states can be neglected in Eq. (30) has to be answered through a comparison with the energy of a quantum of vibration $\hbar\omega$ in the electronic states of interest. It can be shown [20] that the vibronic interactions mixing the electronic states ϵ_m^{el} and ϵ_n^{el} can be neglected if

$$\hbar\omega << |\epsilon_m^{el} - \epsilon_n^{el}|. \tag{37}$$

This condition is the criterion of validity of the adiabatic approximation. It obviously does not hold when $\epsilon_m^{el} = \epsilon_n^{el}$ (degenerate states), or when $\epsilon_m^{el} \approx \epsilon_n^{el}$ (pseudo-degenerate states). This is exactly the case for which the Jahn-Teller theorem was formulated. Therefore, if in the adiabatic approximation we obtain a degenerate ground state of the system, we have to expand the treatment to account for the vibronic mixing of the electronic states, which results in a distortion of the system and lifting of degeneracy.

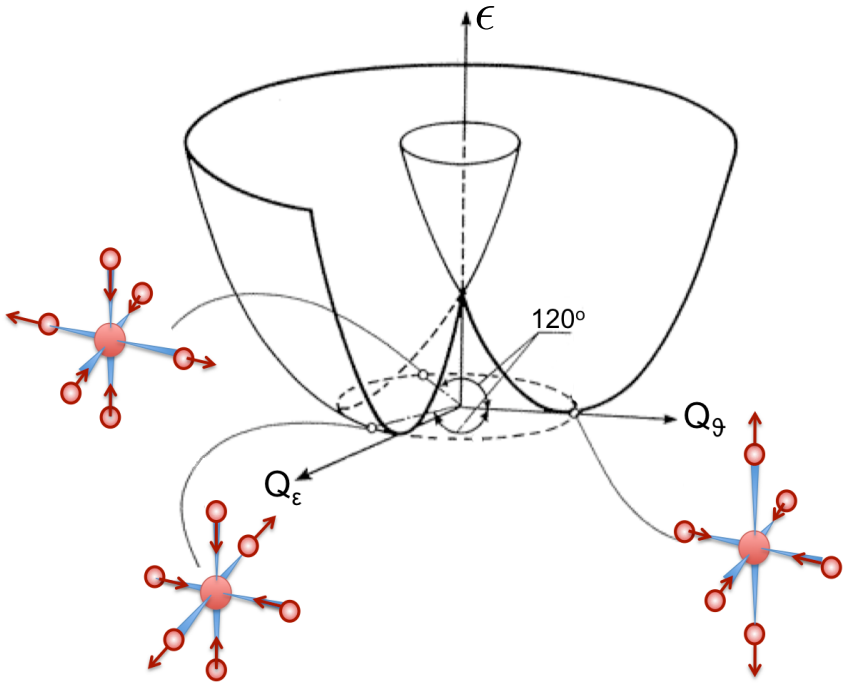


Fig. 9: *Adiabatic potential energy surface (APES), Eq. (40) for the $E \otimes e$ problem. The three indicated atomic configurations correspond to three minima of the APES. In these configurations the ligand octahedrons are tetragonally distorted along the three mutually orthogonal axes. Adopted from Ref. [23].*

Let us assume that the ground state is f-fold degenerate, i.e. the eigenfunctions $\varphi_k,\ k = 1, ..., f$ all have the same energy $\epsilon_k^{el} = \epsilon_0^{el}$. The APES $\epsilon_k(Q),\ k = 1, ...f$ for each of the f states has an additional, vibronic contribution $\epsilon_k^v(Q)$. These contributions are obtained from the prerturbation theory (we consider the nuclear displacements to be small) as the f roots of the secular equation

$$||W_{\gamma,\gamma'} - \epsilon^v|| = 0, \quad \gamma, \gamma' = 1, ..., f, \tag{38}$$

finally yielding the APES in the form

$$\epsilon_k(Q) = \frac{1}{2}\sum_\alpha K_\alpha Q_\alpha^2 + \epsilon_k^v(Q), \quad k = 1, ..., f \tag{39}$$

(compare with Eq. (32)).
As an example, we will return to the case of $LaMnO_3$ and take a look at the solution of what is usually called an $E \otimes e$ problem: interaction of the two-fold degenerate E electronic state with the two-fold degenerate e modes. For the octahedral symmetry under consideration, these are the modes Q_ε and Q_ϑ (see Fig. 8). After solving the secular equation, one obtains the two solutions for the APES that can be written as

$$\epsilon^\pm(\rho, \theta) = \frac{1}{2}K\rho^2 \pm \rho\sqrt{F^2 + G^2\rho^2 + 2FG\rho\cos(3\phi)}, \tag{40}$$

where ρ, θ and ϕ are polar coordinates and the first- and second-order vibronic interactions F and G, respectively, are defined as

$$F = \langle\vartheta|\left(\frac{\partial V}{\partial Q_\vartheta}\right)_0|\vartheta\rangle \quad \text{and} \quad G = \langle\vartheta|\left(\frac{\partial^2 V}{\partial Q_\vartheta \partial Q_\varepsilon}\right)_0|\varepsilon\rangle. \tag{41}$$

The calculated APES is shown in Fig. 9. Due to its specific shape, this surface is often called the "Mexican hat". On the trough of the hat there are three minima, alternating with humps between them. These extrema are defined with

$$\rho^{ext} = \frac{\pm F}{K \mp (-1)^n 2G}, \quad \phi^{ext} = \frac{n\pi}{3}, \quad n = 0, ..., 5. \tag{42}$$

The points $n = 0, 2, 4$ are minima in case $FG < 0$, while $n = 1, 3, 5$ are the saddle points. In case $FG > 0$ it is the other way around. As indicated in Fig. 9, the three minima correspond to tetragonal distortions of the octahedron: an elongation (or compression) along one of the three axes and a compression (or elongation) along the other two. We see that we have obtained a mathematical confirmation of the intuitive picture described at the beginning of this section.

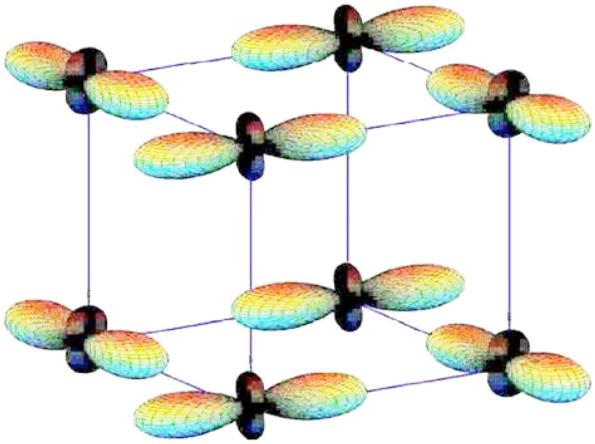


Fig. 10: *Orbital ordering pattern in $LaMnO_3$. Adopted from [22].*

The fact that there are three equivalent minima on the APES does not necessarily mean that the system choses one of them and remains there. There could be several ways of distorting the

structure. However, in crystals it is clear that the distortions need to be compatible with each other. For example, in a perovskite it would be possible to have all the octahedra elongated along the same axis (in a *ferrodistortive* manner), but the direction of the distortions could also be alternating along crystallographic axes (*antiferrodistortive* structure). Since a specific distortion means a specific occupation of the initially degenerate orbitals that caused the distortion, we are talking about the *orbital ordering*. For instance, in Fig. 10 we see the orbital ordering that takes place in $LaMnO_3$ [21]. These are cases of so-called *cooperative Jahn-Teller effect*. The distortion, however, need not be static. There are examples of the compounds, like some mixed-valence systems with Mn^{3+} and Mn^{4+} ions, where the *dynamic Jahn-Teller effect takes place*; the same octahedron is distorted along different axis at different points of time. Under such circumstances, the crystal as a whole seems undistorted. It is also interesting to note that when the APES has minima that are divided by sufficiently small barriers, the quantum-mechanical tunneling effects occur.
A comprehensive overview of Jahn-Teller effects and their occurrence in various crystals and molecules can be found in the book by I. B. Bersuker, Ref. [23].

Acknowledgements

Many thanks to Y. Mokrousov for his help with the figures.

References

[1] I. B. Bersuker, *Electronic structure and properties of transition metal compounds (second edition)*, Wiley, New York, (2010).

[2] W. Kossel, Ann. Phys. (Leipzig), **44**, 229 (1916).

[3] A. Magnus, Z. Anorg. Chem., **124**, 288 (1922).

[4] H. Bethe, Ann. Phys., **3**, 133 (1928).

[5] J. H. Van Vlek, *The theory of electronic and magnetic susceptibilities*, Oxford University Press, London (1965).

[6] L. E. Orgel *An introduction to transition-metal chemistry: ligand field theory*, Wiley, New York, (1960).

[7] C. K. Jorgensen *Modern aspects of ligand field theory*, North-Holland, Amsterdam (1971).

[8] Drawing was made with the VESTA visualization package, K. Momma and F. Izumi, J. Appl. Crystallogr., **41**, 653 (2008).

[9] M. S. Dresselhaus, G. Dresselhaus, A. Jorio, *Group theory: Application to the physics of condensed matter*, Springer-Verlag, Berlin Heidelberg (2008).

[10] R. S. Mulliken, Phys.Rev., **43**, 279 (1933).

[11] P. M. Raccah and J. B. Goodenough, Phys. Rev. **155**, 932 (1967); M. A. Senaris-Rodriguez and J. B. Goodenough, J. Solid State Chem. **116**, 224 (1995).

[12] K. Asai, O. Yokokura, N. Nishimori, H. Chou, J. M. Tranquada, G. Shirane, S. Higuchi, Y. Okajima, and K. Kohn, Phys. Rev. B **50**, 3025 (1994).

[13] M. A. Korotin, S. Yu. Ezhov, I. V. Solovyev, V. I. Anisimov, D. I. Khomskii and G. A. Sawatzky, Phys. Rev. B **54**, 5309 (1996).

[14] C. P. Slichter, Principles of magnetic resonance, Springer-Verlag, Heidelberg (1989).

[15] S. Blundell, *Magnetism in condensed matter*, Oxford University Press, Oxford (2000).

[16] H. A. Kramers, Proc. Amsterdam Acad. **33** 959 (1930).

[17] Jahn, H. A., Teller, E. Proc. R. Soc.,**161**, 220 (1937).

[18] Renner, R. Z. Phys. **92**, 172 (1934).

[19] I. B. Bersuker, Phys. Lett. **20**, 5891 (1966).

[20] I. B. Bersuker, and V. Z. Polinger, *Vibronic interactions in molecules and crystals*, Springer, New York, (1989).

[21] Y. Murakami, J. P. Hill, D. Gibbs, M. Blume, I. Koyama, M. Tanaka, H. Kawata, T. Arima, Y. Tokura, K. Hirota and Y. Endoh, Phys. Rev. Lett. **81**, 582 (1998).

[22] http://www.fkf.mpg.de/keimer/Research/

[23] I. B. Bersuker, *The Jahn-Teller effect*, Cambridge University Press, Cambridge, (2006).

B 6 Dynamics of Macromolecules [1]

R. G. Winkler
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Macromolecule, such as linear polymers, are ubiquitous in everyday products and in biological systems. The spectrum of applications of synthetic polymers is wide and covers such divers fields as medicine (heart valves, blood vessels, drug deliver systems), consumer (containers, clothing, fluid modifiers, suspension stabilizer) and industrial products (automobile parts, adhesives, elastomers, tyres, viscosity modifiers) as well as sports equipment (helmets, balls, golf clubs). (The chemical structures of various polymers are introduced in the contribution E 2 of this lecture notes.) Moreover, macromolecules constitute an integral part of biological systems. Proteins, for example, carry the major part of the mechanical and chemical functions of cells. Desoxyribonucleic acid (DNA) is the molecular basis of heredity and is constructed of a double helix. The extraordinary mechanical and dynamical properties of eukaryotic cells are determined by a three dimensional assembly of protein fibers, the cytoskeleton. A major contribution to this remarkable properties is due to actin filaments, which consist of two strands of globular molecules called actin, and proteins that cross-link them [1, 2]. These molecules are often denoted as semiflexible polymers. A key aspect of DNA structure is its large length-to-width ratio. The diameter of the molecule is approximately 2 nm, whereas typical contour lengths range from micrometers in simple viruses to centimeters in humans. Thus, its structural and dynamical properties can easily by analyzed by optical methods. From a theoretical point of view, this makes DNA an ideal candidate to verify polymer models.

To control the behavior of polymers, an understanding of their dynamical behavior is inevitable. On the one hand, such fundamental knowledge helps to improve production processes, e.g., in injection molding, or to modify the viscoelastic behavior of solutions. On the other hand, by measurements of the polymer dynamics via, e.g., neutron scattering or nuclear magnetic resonance (NMR), insight is gained into the structural properties of polymers and their environment.

Polymers constitute a unique class of material. Thermodynamically, the free energy of condensed matter is given by the Helmholtz free energy $A = E - TS$, where E is the internal energy, T the absolute temperature, and S is the entropy [3–8]. In solids, the internal energy (phonons) dominates over the the entropic part. In contrast, the free energy of polymers is determined by the entropy [9–16]. This implies that other concepts have to be devised to describe the structural and dynamical properties of polymers. Solids often form crystalline lattices, with well-defined equilibrium positions, around which the atoms perform oscillatory motions. In contrast, flexible and semiflexible polymers are able to undergo a huge spectrum of conformational changes at a constant energy. This is illustrated in Fig. 1, where various possible conformations are shown of DNA fragments.

To cope with the huge number of internal degrees of freedom of polymers, their properties are described by statistical physics [9–16]. Here, no longer an individual system is considered, but rather an ensemble of identical systems possessing the same macroscopic properties, e.g., the same number of particles, volume, and mean energy. This provides a distribution function, or a probability density, for their possible conformational states [14, 16]. For the description of the polymer dynamics, stochastic methods are exploited. Here, an effective single polymer approach is adopted, where the influence of the surrounding is taken into account by a stochastic process. In the simplest case, the environment is described by a white-noise random process, i.e., a Gaussian and Markovian process [17, 18]; this applies approximately to polymer melts.

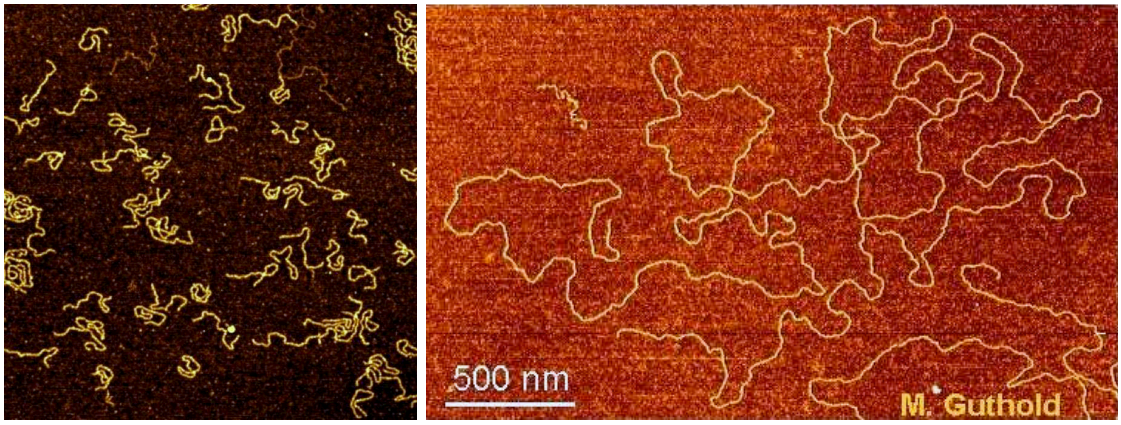


Fig. 1: *AFM images of DNA molecules adsorbed and immobilized on a solid substrate. Short fragments of DNA (left); images size* $1\mu\mathrm{m} \times 1\mu\mathrm{m}$. *Lambda DNA (right), a rather flexible molecule, is composed of 48 502 bp (base pairs), which approximately corresponds a contour length of* $L = 16.49$ *µm. Images are taken from* http://www.dmphotonics.com/AFM-STM-NSOM-sample-quotes/Sample%20quotes%20for%20AFM-STM-NSOM%20Heron.htm *(left) and* http://www.wfu.edu/ gutholdm/research.html *(right).*

For polymers in solution, hydrodynamic interactions between monomers have to be taken into account [14, 19, 20], which are mediated by fluid flow. In melts of very long polymers, the non-crossability of polymers plays a major role and gives rise to dynamical restrictions by neighboring polymers, denoted as entanglements [14]. The polymer dynamics is then described by the reptation model [11, 14, 21, 22].

2 Gaussian Polymer

There are suitable all-atom models for polymers, which are exploited in computer simulations. However, typically only simple, low molecular weight polymers are considered, e.g., alkane molecules. Instead, coarse-grained polymer models are used to study properties of long macromolecules, in particular in analytical theory.

The generic features of a polymer—its conformational degrees of freedom—can be described and captured by a Gaussian model. Typically, this model yields the correct qualitative properties of flexible and semiflexible polymers, often it is even quantitatively correct [14, 23–26]. In such a model, details of the bonds are neglected, even several monomers are combined in an effective segment, merely the connection matters into a linear structure. Denoting the connecting vector between two successive "monomers" along the polymer chain by $\boldsymbol{R}_i = \boldsymbol{r}_i - \boldsymbol{r}_{i-1}$ $(i = 1, \ldots, N)$ (cf. Fig. 2), the Gaussian chain is characterized by the first two moments, namely $\langle \boldsymbol{R}_i \rangle = 0$, i.e., all orientations are equally probable, and $\langle \boldsymbol{R}_i^2 \rangle = l^2$, where $\langle \ldots \rangle$ denotes the ensemble average, at equilibrium. All higher moments can be expressed by these two moments [17]. Considering the non-zero second moment as a constraint, the (configurational) distribution function $\Psi(\{\boldsymbol{r}\})$ and partition function Z can be obtained by the maximum entropy principle, which provides a basis for the systematic derivation of distribution function also in the presence of external fields and bending stiffness [16, 24, 26–28]. The distribution and partition functions for all monomers

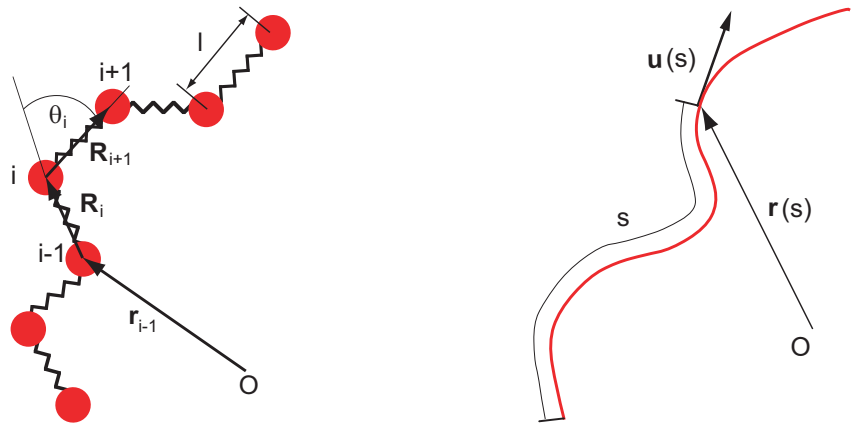


Fig. 2: *Models of a discrete (left) and a continuous (right) linear polymer. The tangent vector* $\boldsymbol{u}(s) = \partial \boldsymbol{r}(s)/\partial s$ *is obtained in the limit* $\boldsymbol{R}_i/l \to \boldsymbol{u}(s)$, $il \to s$ *for* $l \to 0$.

are given by the Gaussian

$$\Psi(\{\boldsymbol{r}\}) = \frac{1}{Z} \exp\left(-\frac{3}{2l^2}\sum_{i=1}^{N} \boldsymbol{R}_i^2\right), \tag{1}$$

$$Z = \int \exp\left(-\frac{3}{2l^2}\sum_{i=1}^{N} \boldsymbol{R}_i^2\right) d^{3N}R, \tag{2}$$

where $\boldsymbol{r}_0 = 0$ is used to remove the translation degree of freedom. Note that the total partition function $Z_{tot} = ZZ_k$ is the product of the configurational part (2) and the contribution by the kinetic energy $Z_k = (2\pi m k_B T)^{3N/2}$ ($\boldsymbol{v}_0 = 0$), where m is the mass of a monomer. Equation (1) and (2) are reminiscent of a system of independent harmonic oscillators with the potential energy $V_i = 3k_BT\boldsymbol{R}_i^2/(2l^2)$. From the point of view of the underlying physics, however, the equations capture the conformational degrees of freedom of a polymer rather then energetic aspects. The Gaussian polymer is self-similar, as the distribution of segment lengths is similar to the end-to-end vector distribution function

$$\Psi(\boldsymbol{r}_N) = \left(\frac{3}{2\pi N l^2}\right) \exp\left(-\frac{3}{2Nl^2}\boldsymbol{r}_N^2\right), \tag{3}$$

with the mean square end-to-end distance $\langle(\boldsymbol{r}_N - \boldsymbol{r}_0)^2\rangle = Nl^2$. The maximum entropy principle yields the conformational entropy [16, 28]

$$S = k_B \ln Z + \frac{3}{2} N k_B, \tag{4}$$

where the last term accounts for the potential energy of the bonds.

For various calculations, a continuous polymer model is advantageous. From a physical point of few, it does not matter whether a continuous or discrete model is used. Because of the adopted abstraction of a real polymer, the chosen description applies on large length scales only. For a Gaussian polymer model, only properties are described on the length scale larger than the segment length l.

The continuum limit $N \to \infty$, $l \to 0$, such that $L = Nl = const.$, where L is the contour length of the polymer, yields

$$Z = \int \exp\left(-\sum_{i=1}^{N} \frac{3}{2l^2}\boldsymbol{R}_i^2\right) d^{3N}r \overset{l\to 0, N\to\infty}{\longrightarrow} \int \exp\left(-\frac{3}{2l}\int_0^L \left(\frac{\partial \boldsymbol{r}}{\partial s}\right)^2 ds\right) \mathcal{D}^3 r,$$

i.e., a functional integral (path integral) is obtained. However, in the limit $l \to 0$, $3/(2l) \to \infty$ and the partition function approaches zero ($\lim_{l\to 0} Z = 0$), i.e., the polymer collapses into a point. The reason is that the distribution function (1) corresponds to a Wiener process [17], and the derivative of the 'trajectory', i.e., the polymer contour, does not exist. A meaningful continuum limit can only be performed for a semiflexible polymer [16, 24]. For the Gaussian polymer, a continuum description can only be adopted on length scales $s \gg l$. Hence, the length l is assumed to be small, but finite and the partition function reads

$$Z = \int \exp\left(-\frac{3}{2l}\int_0^L \left(\frac{\partial \boldsymbol{r}}{\partial s}\right)^2 ds\right) \mathcal{D}^3 r. \tag{5}$$

l is sometimes denoted as Kuhn length l_K and is related to the polymer persistence length l_p via $l_K = l = 2l_p$.
The distribution function of two points $\boldsymbol{r}(s)$, $\boldsymbol{r}(s')$, with $\Delta\boldsymbol{r} = \boldsymbol{r}(s) - \boldsymbol{r}(s')$, is

$$\Psi(\Delta\boldsymbol{r}) = \left(\frac{3}{2\pi\sigma^2(s,s')}\right)^{3/2} \exp\left(-\frac{3\Delta\boldsymbol{r}^2}{2\sigma^2(s,s')}\right), \tag{6}$$

where $\sigma^2(s,s') = \left\langle \left(\boldsymbol{r}(s) - \boldsymbol{r}(s')\right)^2 \right\rangle = l|s-s'|$. The radius of gyration is obtained as $R_G^2 = lL/6$.

In order to describe the properties of semiflexible polymers such as DNA or actin filaments, specifically their resistance to bending, a bending energy has to be taken into account. Kratky and Porod suggested an energy, which is determined by the square of the local curvature of the continuous polymer model [29]

$$\frac{V_B}{k_B T} = \frac{\epsilon_p}{2}\int_0^L \frac{1}{R(s)^2} ds = \frac{\epsilon_p}{2}\int_0^L \left(\frac{\partial^2 \boldsymbol{r}(s)}{\partial s^2}\right)^2 ds. \tag{7}$$

Taking bending restrictions into account, the partition function of a Gaussian semiflexible polymer is [16, 20, 23–27, 30]

$$Z = \int \exp\left(-\frac{3}{4l_p}\int_0^L \left(\frac{\partial \boldsymbol{r}(s)}{\partial s}\right)^2 ds - \frac{3}{4}\left[\left(\frac{\partial \boldsymbol{r}(0)}{\partial s}\right)^2 + \left(\frac{\partial \boldsymbol{r}(L)}{\partial s}\right)^2\right] - \frac{3l_p}{4}\int_0^L \left(\frac{\partial^2 \boldsymbol{r}(s)}{\partial s^2}\right)^2 ds\right) \mathcal{D}^3 r\,. \tag{8}$$

Since this is a Gaussian function, Eq. (6) applies with

$$\sigma^2(s,s') = 2l_p|s-s'| - 2l_p^2\left(1 - e^{-|s-s'|/l_p}\right). \tag{9}$$

This equation reduces to the expression $\sigma^2 = 2l_p|s-s'|$ for $L/l_p \gg |s-s'|/l_p \gg 1$, i.e., flexible polymer behavior, and $\sigma^2 = (s-s')^2$ for $|s-s'|/l_p \ll 1$, which corresponds to (local) rodlike behavior. In the case $L/l_p \ll 1$, the latter applies to all distances $|s-s'|$.

3 Dynamics of Flexible Polymers

The dynamics of a polymer is determined by its interactions with the environment. For a polymer in a melt, i.e., in a dense environment of similar polymers, the local frictional interactions dominate. In contrast, the dynamics of a polymer in dilute solution is determined by fluid mediated interactions. Here, the motion of a monomer implies fluid motion due to no-slip boundary conditions, i.e., locally the fluid at the position of a particle moves with the same velocity as the particle itself and vice versa. This fluid motion affects the dynamics of other particles and is denoted as hydrodynamic interactions [14, 19]. Aside from that, intramolecular forces are present. In the continuum limit, the "potential energy"

$$V = \frac{3k_BT}{2l}\int_0^L \left(\frac{\partial \boldsymbol{r}(s)}{\partial s}\right)^2 ds \tag{10}$$

yields the local force

$$\boldsymbol{F}(s) = \frac{3k_BT}{l}\frac{\partial^2 \boldsymbol{r}(s)}{\partial s^2} \tag{11}$$

on the point $\boldsymbol{r}$ at s along the polymer contour. Note, all points of the polymer are free.

In the following, only the limit $L/l \gg 1$ will be considered, i.e., only the dynamics of flexible polymers will be addressed in detail. The dynamics of semiflexible polymers is discussed in Refs. [20, 23, 25–27].

3.1 Free-Draining Dynamics—Rouse Model

The friction-dominated dynamics, i.e., no hydrodynamic interactions, is denoted as free-draining dynamics (Rouse model [14, 31]). Here, the equation of motion of the point $\boldsymbol{r}(s,t)$ is given by

$$\gamma\frac{\partial \boldsymbol{r}(s,t)}{\partial t} = \frac{3k_BT}{l}\frac{\partial^2 \boldsymbol{r}(s,t)}{\partial s^2} + \boldsymbol{\Gamma}(s,t), \tag{12}$$

with the friction constant γ and the stochastic force $\boldsymbol{\Gamma}(s,t)$. Since a polymer is long and hence its dynamics is slow compared to fluctuations in the environment, the over-damped equation is considered only, i.e., inertia effects are neglected—their influence on the dynamics has decayed on the polymer-relevant time scales (see also contribution B3) [14, 19]. The equation of motion is complemented by the boundary conditions

$$\left.\frac{\partial \boldsymbol{r}(s,t)}{\partial s}\right|_{s=0,L} = 0 \tag{13}$$

and the correlation functions of the stochastic force, which is assumed to be Gaussian and Markovian (white noise),

$$\langle \boldsymbol{\Gamma}(s,t)\rangle = 0\ , \tag{14}$$

$$\langle \Gamma_\alpha(s,t)\Gamma_{\alpha'}(s',t')\rangle = 2\gamma k_BT\delta_{\alpha\alpha'}\delta(t-t')\delta(s-s')\ . \tag{15}$$

Equation (12) is a linear partial differential equation, which is easily solved by an eigenfunction expansion. The relevant eigenvalue equation is

$$\frac{3k_BT}{l}\frac{\partial^2}{\partial s^2}\varphi_n(s) + \xi_n\varphi_n = 0\ , \tag{16}$$

with the eigenfunctions φ_n and the eigenvalues ξ_n. With the boundary conditions $\partial\varphi_n/\partial s = 0$ at $s = 0, L$ follows

$$\varphi_0 = \sqrt{\frac{1}{L}}\,, \qquad \xi_0 = 0\,, \tag{17}$$

$$\varphi_n = \sqrt{\frac{2}{L}}\cos\frac{n\pi s}{L}\,, \qquad \xi_n = \frac{3\pi^2 k_B T n^2}{lL^2} \quad \forall\, n > 0\,. \tag{18}$$

With the eigenfunction expansion

$$\boldsymbol{r}(s,t) = \sum_{n=0}^{\infty} \boldsymbol{\chi}_n(t)\varphi_n(s)\,, \qquad \boldsymbol{\Gamma}(s,t) = \sum_{n=0}^{\infty} \boldsymbol{\Gamma}_n(t)\varphi_n(s), \tag{19}$$

Eq. (12) yields the Langevin equations

$$\frac{d}{dt}\boldsymbol{\chi}_0 = \frac{1}{\gamma}\boldsymbol{\Gamma}_0\,; \qquad \frac{d}{dt}\boldsymbol{\chi}_n = -\frac{1}{\tau_n}\boldsymbol{\chi}_n + \frac{1}{\gamma}\boldsymbol{\Gamma}_n\,,\ n > 0\,, \tag{20}$$

for the normal-mode amplitudes $\boldsymbol{\chi}_n$, with the relaxation times

$$\tau_n = \frac{\gamma}{\xi_n} = \frac{\gamma l L^2}{3\pi^2 k_B T n^2} \equiv \frac{\tau_R}{n^2}\,, \quad \forall n > 0, \tag{21}$$

and the correlations $\left\langle \Gamma_n^{\alpha}(t)\Gamma_m^{\alpha'}(t')\right\rangle = 2\gamma k_B T\delta_{\alpha\alpha'}\delta_{nm}\delta(t-t')$. $\tau_R = \gamma l L^2/(3\pi^2 k_B T)$ is the Rouse relaxation time [14]. It is important to note that the relaxation times exhibit a quadratic length and mode-number dependence. The solution of the Langevin equation (20) (Ornstein-Uhlenbeck process) is [17]

$$\boldsymbol{\chi}_0(t) = \boldsymbol{\chi}_0(0) + \frac{1}{\gamma}\int_0^t \boldsymbol{\Gamma}_0(t')\,dt'\,; \qquad \boldsymbol{\chi}_n(t) = \frac{1}{\gamma}\int_{-\infty}^t \exp\left(-\frac{t-t'}{\tau_n}\right)\boldsymbol{\Gamma}_n(t')\,dt'\,. \tag{22}$$

Hence, the solution of Eq. (12) is given by

$$\boldsymbol{r}(s,t) = \boldsymbol{\chi}_0(t)\varphi_0 + \sum_{n=1}^{\infty}\frac{1}{\gamma}\int_{-\infty}^t \exp\left(-\frac{t-t'}{\tau_n}\right)\boldsymbol{\Gamma}_n(t')dt'\varphi_n(s). \tag{23}$$

With that expression, various dynamical properties of polymers can be calculated.

(i) **Center-of-mass mean square displacement**

The center-of-mass of a polymer is defined as

$$\boldsymbol{r}_{cm}(t) = \frac{1}{L}\int_0^L \boldsymbol{r}(s,t)\,ds = \boldsymbol{\chi}_0(t)\varphi_0. \tag{24}$$

Hence, its mean square displacement (MSD) is

$$\left\langle \Delta\boldsymbol{r}_{cm}(t)^2\right\rangle = \frac{1}{L}\left\langle(\boldsymbol{\chi}_0(t) - \boldsymbol{\chi}_0(0))^2\right\rangle = \frac{1}{\gamma^2 L}\int_0^t\int_0^t \langle\boldsymbol{\Gamma}_0(t')\boldsymbol{\Gamma}_0(t'')\rangle\,dt'dt'' = \frac{6k_BT}{\gamma L}t, \tag{25}$$

or $\left\langle \Delta\boldsymbol{r}_{cm}(t)^2\right\rangle = \left\langle(\boldsymbol{r}_{cm}(t) - \boldsymbol{r}_{cm}(0))^2\right\rangle = 6D_R t$, where $D_R = k_BT/(\gamma L)$ is the center-of-mass diffusion coefficient, which inversely proportional to the polymer length.

(ii) **Segmental mean square displacement**

The mean square displacement of the point $\boldsymbol{r}(s,t)$ is given by

$$\left\langle \Delta\boldsymbol{r}(s,t)^2 \right\rangle = \left\langle [\boldsymbol{r}(s,t)-\boldsymbol{r}(s,0)]^2 \right\rangle = \left\langle \Delta\boldsymbol{r}_{cm}(t)^2 \right\rangle + \sum_{n=1}^{\infty} 2\varphi_n(s)^2 \left\langle \boldsymbol{\chi}_n^2 \right\rangle \left(1-e^{-t/\tau_n}\right).$$

With the correlation functions ($n,m>0$)

$$\begin{aligned}\langle \boldsymbol{\chi}_n(t)\boldsymbol{\chi}_m(0)\rangle &= \frac{1}{\gamma^2}\int_{-\infty}^{t}\int_{-\infty}^{0} \exp\left(-\frac{t-t'}{\tau_n}\right)\exp\left(\frac{t''}{\tau_m}\right)\langle \boldsymbol{\Gamma}_n(t')\boldsymbol{\Gamma}_m(t'')\rangle\, dt'dt'' \qquad (26)\\ &= \frac{6k_BT}{\gamma}\delta_{nm}e^{-t/\tau_n}\int_{-\infty}^{0} e^{2t'/\tau_n}dt' = \frac{3k_BT}{\gamma}\tau_n\delta_{nm}e^{-t/\tau_n} = \left\langle \boldsymbol{\chi}_n^2\right\rangle \delta_{nm}e^{-t/\tau_n}\end{aligned}$$

follows

$$\left\langle \Delta\boldsymbol{r}(s,t)^2 \right\rangle = \left\langle \Delta\boldsymbol{r}_{cm}(t)^2 \right\rangle + \frac{6k_BT}{\gamma}\sum_{n=1}^{\infty}\varphi_n(s)^2\tau_n\left(1-e^{-t/\tau_n}\right). \qquad (27)$$

For $t/\tau_R \gg 1$, the second term is a constant, because $e^{-t/\tau_n} \ll 1$, $\forall n>0$. Since $\langle \Delta\boldsymbol{r}_{cm}(t)^2\rangle \sim t$, the segmental MSD is equal to the center-of-mass MSD in the long-time limit. However, for times $t/\tau_R \ll 1$, the sum is dominated by large mode-number contributions. Then, the MSDs of the various parts of the polymer are different. For example, the center of the polymer ($s=L/2$) moves by a factor of two slower than the ends ($s=0,L$). Averaging over the polymer contour: $\langle \Delta\boldsymbol{r}(t)^2\rangle = \int \langle \Delta\boldsymbol{r}(s,t)^2\rangle\, ds/L$, yields

$$\left\langle \Delta\boldsymbol{r}(t)^2 \right\rangle = \left\langle \Delta\boldsymbol{r}_{cm}(t)^2 \right\rangle + \frac{6k_BT}{\gamma L}\sum_{n=1}^{\infty}\tau_n\left(1-e^{-t/\tau_n}\right). \qquad (28)$$

For $t/\tau_R \ll 1$, the difference between $x_{n+1}=(n+1)\sqrt{t/\tau_R}$ and $x_n = n\sqrt{t/\tau_R}$ is very small and the sum can be replaced by an integral over $x_n = n\sqrt{t/\tau_R}$, i.e.,

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\tau_R}{n^2}\left[1-e^{-tn^2/\tau_R}\right] &= \sqrt{t\tau_R}\sum_{n=1}^{\infty}\frac{\tau_R}{n^2t}\left[1-e^{-tn^2/\tau_R}\right]\Delta n\sqrt{\frac{t}{\tau_R}}\\ &\longrightarrow \sqrt{t\tau_R}\int_0^{\infty}\frac{1}{x^2}(1-e^{-x^2})dx = \sqrt{t\tau_R}\sqrt{\pi}.\end{aligned}$$

Hence,

$$\left\langle \Delta\boldsymbol{r}(t)^2 \right\rangle = \left\langle \Delta\boldsymbol{r}_{cm}(t)^2 \right\rangle + \frac{6k_BT}{\gamma L}\sqrt{\pi t\tau_R} \overset{t/\tau_R\ll 1}{\longrightarrow} \frac{2lL}{\sqrt{\pi^3}}\sqrt{\frac{t}{\tau_R}}, \qquad (29)$$

i.e. the segments display an anomalous (fractal) diffusion behavior.
The center-of-mass (25) and segmental (28) mean square displacement are displayed in Fig. 3. For $t/\tau_R \ll 1$, the total MSD is dominated by the internal dynamics, i.e., the second term in Eq. (28). As is indicated by the agreement between the full numerical result and the analytical approximation, in that time regime the MSD increase as $t^{1/2}$. For $t/\tau_R \gg 1$, the MSD is given by the center-of-mass displacement.

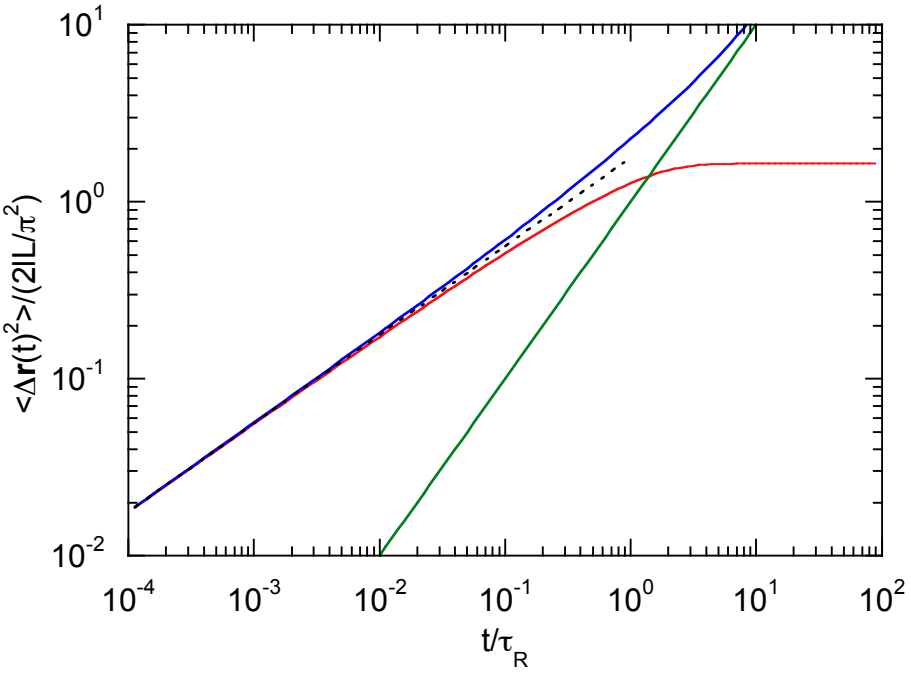


Fig. 3: *Mean square displacements of a free-draining polymer. The green line indicates* $\langle \Delta \boldsymbol{r}_{cm}(t)^2 \rangle$*, the blue line the total MSD* $\langle \Delta \boldsymbol{r}(t)^2 \rangle$*, and the red line the segmental MSD in the center-of-mass reference frame:* $\langle \Delta \boldsymbol{r}(t)^2 \rangle - \langle \Delta \boldsymbol{r}_{cm}(t)^2 \rangle$*. The red line approaches the asymptotic value* $2R_G^2$ *in the limit* $t \to \infty$*. The dotted line represents the analytical approximation (29).*

3.2 Non-Draining Dynamics—Zimm Model

In dilute solution, the polymer dynamics is governed by hydrodynamic interactions. To determine the velocity of a particle at $\boldsymbol{r}(s)$, the fluid velocity field $\boldsymbol{v}(\boldsymbol{r})$ is needed at that position. Since the polymer motion is intimately coupled to the motion of the background fluid, aside from the dynamical equation of the polymer, a dynamical equation for the fluid is needed. To tackle this problem analytically, the fluid is considered a continuum and its dynamics obeys the Navier-Stokes equation or, more precisely, the Stokes equation, since the low Reynolds number limit is considered [14, 19]. As solution of the Stokes equation in response to a volume force $\boldsymbol{f}$, the velocity field is obtained

$$\boldsymbol{v}(\boldsymbol{r}) = \int \boldsymbol{\Omega}(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{f}(\boldsymbol{r}') \, d^3 r', \tag{30}$$

with the *Oseen* tensor

$$\boldsymbol{\Omega}(\boldsymbol{r}) = \frac{1}{8\pi\eta r} \left(\mathbf{I} + \frac{\boldsymbol{r} : \boldsymbol{r}}{r^2} \right) , \tag{31}$$

where η is the solvent viscosity. A detailed derivation is provide in Appendix A. The Oseen tensor decays like $1/r$. Hence, hydrodynamic interactions are of long-range nature and cannot be neglected. For the continuous polymer chain, the force density is

$$\boldsymbol{f}(\boldsymbol{r}) = \int \left[\boldsymbol{F}(s) + \boldsymbol{\Gamma}(s) \right] \delta(\boldsymbol{r} - \boldsymbol{r}(s)) \, ds, \tag{32}$$

thus,

$$\boldsymbol{v}(\boldsymbol{r}) = \int \boldsymbol{\Omega}(\boldsymbol{r} - \boldsymbol{r}(s)) \left[\boldsymbol{F}(s) + \boldsymbol{\Gamma}(s)\right] ds. \tag{33}$$

In the presence of a solvent, the friction between a particle at $\boldsymbol{r}(s)$ and the fluid background is described by

$$\gamma \left[\dot{\boldsymbol{r}}(s,t) - \boldsymbol{v}(\boldsymbol{r}(s,t))\right] = \boldsymbol{F}(s,t) + \boldsymbol{\Gamma}(s,t). \tag{34}$$

For a particle instantaneously following the fluid flow field $\dot{\boldsymbol{r}}(s,t) = \boldsymbol{v}(\boldsymbol{r}(s,t))$ ($\gamma \to \infty$). In case of a finite slip

$$\dot{\boldsymbol{r}}(s,t) = \boldsymbol{v}(\boldsymbol{r}(s,t)) + \frac{1}{\gamma}\left[\boldsymbol{F}(s,t) + \boldsymbol{\Gamma}(s,t)\right],$$

and the equation of motion of the polymer becomes with Eqs. (11), (33)

$$\frac{\partial \boldsymbol{r}(s,t)}{\partial t} = \int_0^L \mathbf{H}(\boldsymbol{r}(s) - \boldsymbol{r}(s')) \left[\frac{3k_BT}{l}\frac{\partial^2 \boldsymbol{r}(s',t)}{\partial s'^2} + \boldsymbol{\Gamma}(s',t)\right] ds', \tag{35}$$

where the hydrodynamic tensor $\mathbf{H}(\boldsymbol{r}(s))$ is defined as

$$\mathbf{H}(\boldsymbol{r}(s) - \boldsymbol{r}(s')) = \frac{\delta(s-s')}{\gamma}\mathbf{I} + \boldsymbol{\Omega}(\boldsymbol{r}(s) - \boldsymbol{r}(s')). \tag{36}$$

Hydrodynamic interactions control the dynamics of the polymer in solution, however, the polymer conformational properties are not affected. The way conformational space is sampled is different for a non-draining and a free-draining polymer, but not the conformational space itself, at least at equilibrium. This can easily be shown by the stationary state solution of the Fokker-Planck equation corresponding to the Langevin equation (35).

Equation (35) is non-linear and therefore difficult to solve. Thus, for flexible polymers typically the preaveraging approximation is adopted, as originally proposed by Zimm [14, 32]. Here, the hydrodynamic tensor is replaced by its spatially averaged expression $\boldsymbol{\Omega}(\boldsymbol{r}(s) - \boldsymbol{r}(s')) \longrightarrow \langle\boldsymbol{\Omega}(\boldsymbol{r}(s) - \boldsymbol{r}(s'))\rangle$. Utilizing the Gaussian joint probability density (6), the Oseen tensor becomes

$$\langle\boldsymbol{\Omega}(\boldsymbol{r}(s) - \boldsymbol{r}(s'))\rangle = \frac{\mathbf{I}}{6\pi\eta}\left\langle\frac{1}{|\boldsymbol{r}(s) - \boldsymbol{r}(s')|}\right\rangle = \frac{\mathbf{I}}{3\pi\eta}\sqrt{\frac{3}{2\pi l|s-s'|}} \equiv \Omega(s-s')\mathbf{I}. \tag{37}$$

This expression is isotropic, because any non-diagonal part of $\boldsymbol{\Omega}$ vanishes (space is homogeneous and isotropic). Hence, $\mathbf{H}(s-s') = \mathbf{I}\mathrm{H}(\mathrm{s}-\mathrm{s}') = \mathbf{I}\left[\delta(s-s')/3\pi\eta + \Omega(s-s')\right]$, where $\gamma = 3\pi\eta$ has be used. For the latter, it is assumed that every segment is a sphere of diameter l with the friction coefficient $\hat{\gamma} = 3\pi\eta l$ (Stokes law). This leads to the friction density $\hat{\gamma}/l \to \gamma = 3\pi\eta$ in the continuum limit $l \to 0$.

Within the preaveraging approximation, the equation of motion (35) is linear and the eigenfunction expansion (19) leads to the equations of motion for the mode amplitudes

$$3\pi\eta\frac{\partial \boldsymbol{\chi}_k}{\partial t} = \sum_{n=0}^{\infty} \mathrm{H}_{kn}\left(-\frac{3\pi\eta}{\tau_n}\boldsymbol{\chi}_n + \boldsymbol{\Gamma}_n\right), \tag{38}$$

where $\mathrm{H}_{kn} = 3\pi\eta\Omega_{kn} + \delta_{kn}$ and

$$\Omega_{kn} = \frac{1}{\sqrt{6\pi^3}\eta} \int_0^L \int_0^L \varphi_k(s') \frac{1}{\sqrt{l|s-s'|}} \varphi_n(s) ds ds'. \tag{39}$$

The correlation functions of the random forces are given by

$$\langle \Gamma_{n\alpha}(t) \Gamma_{k\alpha'}(t') \rangle = 6\pi\eta k_B T \delta_{\alpha\alpha'} \delta(t-t') \mathrm{H}^{-1}_{nk}. \tag{40}$$

Eq. (38) is an infinite system of coupled linear differential equations, which can be solved (numerically) by diagonalization. For an analytical solution, the approximation

$$\Omega_{kn} \approx \sqrt{\frac{2}{3\pi^3}} \frac{\delta_{kn}}{\eta L} \int_0^L \frac{L-s}{\sqrt{ls}} \cos\left(\frac{n\pi s}{L}\right) ds \tag{41}$$

is used [14]. A numerical calculation confirms that the matrix Ω_{nk} is almost diagonal for the considered flexible polymers ($L/l \gg 1$). The equations of motion for the normal-mode amplitudes are then given by

$$\frac{\partial \boldsymbol{\chi}_n}{\partial t} = -\frac{\mathrm{H}_{nn}}{\tau_n} \boldsymbol{\chi}_n + \frac{\mathrm{H}_{nn}}{3\pi\eta} \boldsymbol{\Gamma}_n = -\frac{1}{\tilde{\tau}_n} \boldsymbol{\chi}_n + \frac{\mathrm{H}_{nn}}{3\pi\eta} \boldsymbol{\Gamma}_n, \tag{42}$$

with the relaxation times

$$\tilde{\tau}_n = \frac{\tau_n}{1 + 3\pi\eta\Omega_{nn}} \approx \frac{\eta}{\sqrt{3\pi} k_B T} \left(\frac{lL}{n}\right)^{3/2} \equiv \frac{\tau_Z}{n^{3/2}} \tag{43}$$

and the Zimm time $\tau_Z = \eta (lL)^{3/2} / (\sqrt{3\pi} k_B T)$ [14, 32]. The relaxation times τ_n are given in Eq. (21). Hydrodynamic interactions lead to a weaker dependence of the relaxation times on polymer length and mode number as the Rouse model. As for the non-hydrodynamic case, the normal-mode amplitudes are given by

$$\boldsymbol{\chi}_n(t) = \frac{\tau_n}{3\pi\eta\tilde{\tau}_n} \int_{-\infty}^{t} \exp\left(-\frac{t-t'}{\tilde{\tau}_n}\right) \boldsymbol{\Gamma}_n(t') dt', \tag{44}$$

and the correlation functions by

$$\langle \boldsymbol{\chi}_k(t) \boldsymbol{\chi}_n(0) \rangle = \frac{k_B T}{\pi\eta} \tau_n \delta_{kn} e^{-t/\tilde{\tau}_n} \ , \ \forall k, n \neq 0, \tag{45}$$

which include the relaxation times $\tilde{\tau}_n$. Hence, the correlation functions decay exponentially and are universal functions of $t/\tilde{\tau}_n$.

We performed mesoscale hydrodynamic simulations of a discrete Gaussian polymer model [33, 34]. In particular, we calculated the correlation functions of the normal-mode amplitudes. For a discrete model of N monomers, the amplitudes are given by $\boldsymbol{\chi}_n = \sum_{i=1}^{N} \boldsymbol{r}_i \cos(n\pi[i - 1/2]/N)$ and the non-draining relaxation times by $\tilde{\tau}_n \sim (n/N)^{1/2} / \sin^2(n\pi/N)$, which turn into $\tilde{\tau}_n \sim (L/n)^{3/2}$ [Eq. 43] in the continuum limit. Figure 4 displays autocorrelation functions for the mode amplitudes. Within the accuracy of the simulations, the correlation functions decay exponentially and exhibit the scaling behavior predicted by the Zimm model.

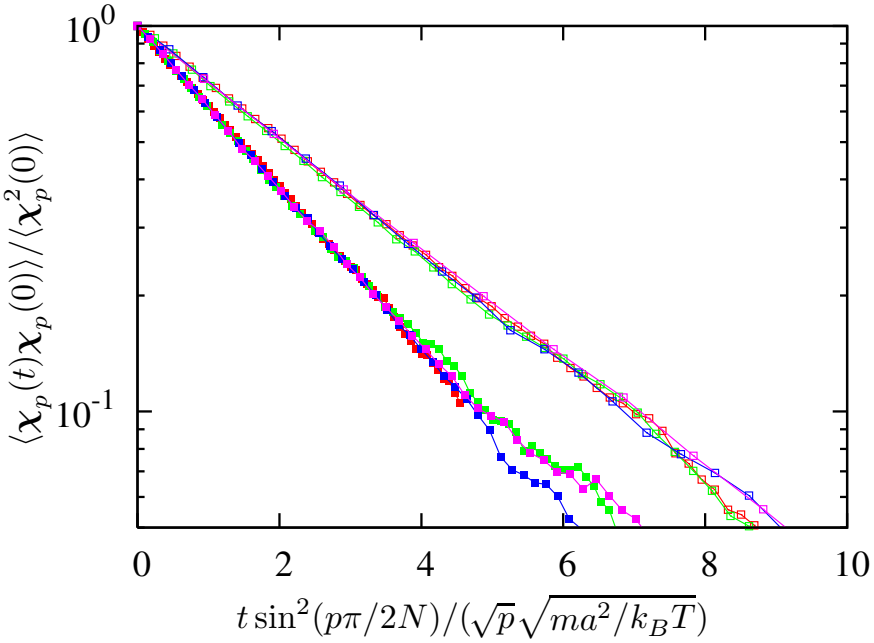


Fig. 4: *Correlation functions of the mode amplitudes of a non-draining Gaussian polymer for the modes $p = 1 - 4$. The polymer lengths are $N = 20$ (right) and $N = 40$ (left) [34].*

Experimentally, the normal-mode theory has been tested by measuring the fluctuations of single DNA molecules held in a partially extended state by optical tweezers [35]. The authors find that the motion of DNA can be described by linearly independent normal modes. However, they also found that the dependence of the relaxation times on the mode number is a function of the DNA extension a [23]. Figure 5 compares experimental data with theoretical results obtained by a semiflexible polymer model (see Eq. (8)) [23]. The analytical solution suggests that hydrodynamic interactions, the finite extensibility of the polymer, and its stiffness determine the observed relaxation time dependence. A free-draining, flexible polymer model does not describe the measured behavior.

(i) **Center-of-mass mean square displacement**

Equation (42) gives

$$\boldsymbol{\chi}_0(t) = \boldsymbol{\chi}_0(0) + \frac{\mathrm{H}_{00}}{3\pi\eta}\int_0^t \boldsymbol{\Gamma}_0(t')dt', \tag{46}$$

which yields the center-of-mass mean square displacement

$$\left\langle \Delta \boldsymbol{r}_{cm}(t)^2 \right\rangle = \frac{\mathrm{H}_{00}^2}{(3\pi\eta)^2}\int_0^t\int_0^t \langle \boldsymbol{\Gamma}_0(t')\boldsymbol{\Gamma}_0(t'')\rangle\, dt'dt'' = \frac{6k_BT}{3\pi\eta L}\mathrm{H}_{00}t, \tag{47}$$

and the diffusion coefficient ($L/l \gg 1$)

$$D_Z = \frac{k_BT}{3\pi\eta L}\mathrm{H}_{00} = \frac{8k_BT}{3\sqrt{6\pi^3}\eta}\frac{1}{\sqrt{lL}}. \tag{48}$$

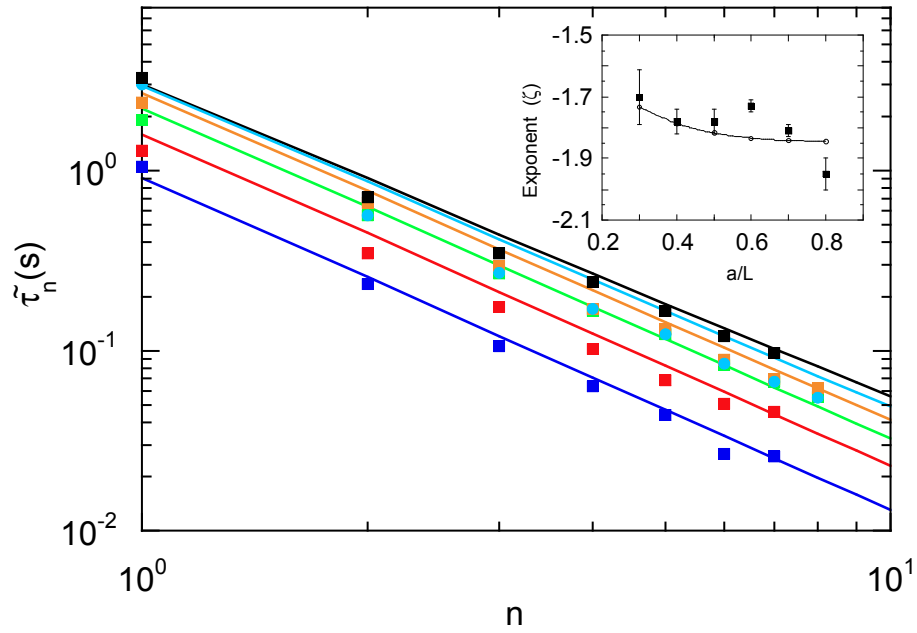


Fig. 5: *Relaxation times of DNA molecules for the extensions* $a/L = 0.3$, 0.4, 0.5, 0.6, 0.7, *and* 0.8 *(top to bottom) [23]. The inset shows the exponents* ζ *of the relation* $\tilde{\tau}_n \sim n^\zeta$ *corresponding to the various curves. The experimental data (squares) are taken from Ref. [35].*

Evidently, diffusion in a dilute solution is faster than in a polymer melt at the same viscosity, as expressed by the ration $D_Z/D_R \sim \sqrt{L/l}$. The long-range hydrodynamic interactions enhance the diffusive dynamics by collective motion of the fluid.

Experimentally, the molecular weight dependence $D_Z \sim L^{-1/2}$ has been confirmed for polymers under Θ conditions, where excluded-volume interactions can be neglected, i.e., conditions under which the described model applies [14, 36–39]. However, the experimental value D_Z is about 15% smaller than the theoretical value (48) [14, 40]. Similar, computer simulations of polymers in solution yield the same molecular weight dependence and even D_Z is in close agreement with the theoretical prediction [34, 41, 42]. This displays the remarkable success of the Zimm model.

(ii) **Segmental mean square displacement**

Similar to the free-draining case, the averaged segmental mean square displacement is given by

$$\left\langle \Delta \boldsymbol{r}(t)^2 \right\rangle = 6D_Z t + \frac{2k_B T}{\pi \eta L} \sum_{n=1}^{\infty} \tau_n \left(1 - e^{-t/\tilde{\tau}_n}\right). \tag{49}$$

With the Zimm relaxation times (43), the sum is replaced by an integral as follows ($x_n = n(t/\tau_z)^{2/3}$)

$$\sum_{n=1}^{\infty} \frac{\tau_R}{n^2} \left[1 - e^{-tn^{3/2}/\tau_z}\right] \to \left(\frac{t}{\tau_z}\right)^{2/3} \tau_R \int_0^\infty \frac{1}{x^2} \left[1 - e^{-x^{3/2}}\right] dx. \tag{50}$$

Hence, $\langle \Delta \boldsymbol{r}^2 \rangle \sim t^{2/3}$, i.e., hydrodynamic interactions lead to a stronger time dependence of the segmental MSD compared to the free-draining case.

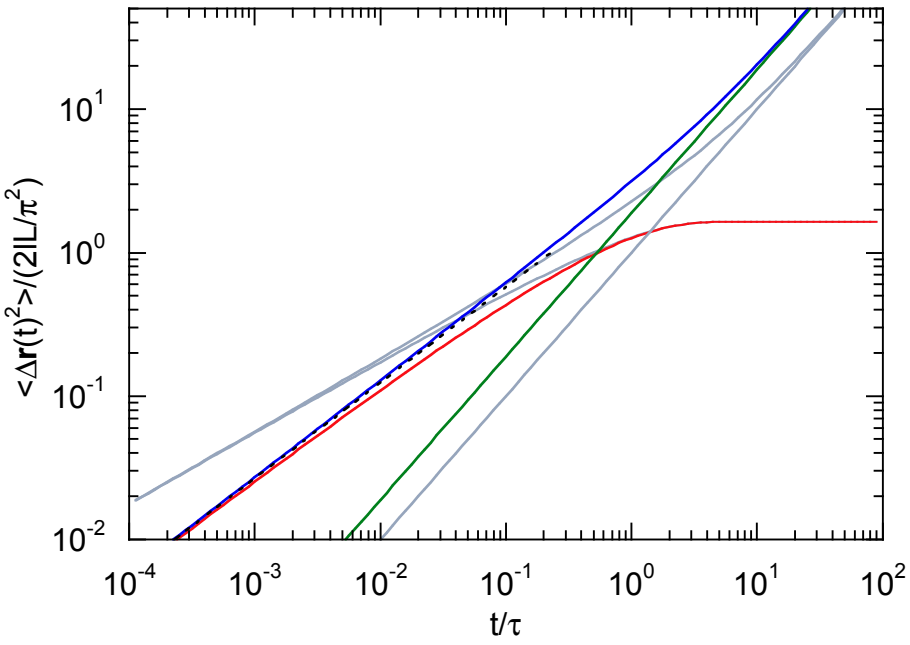


Fig. 6: *Mean square displacements of polymers in dilute solution. The gray lines correspond to the Rouse model (see Fig. 3) and the colored lines to the non-draining Zimm model. The green line indicates* $\langle \Delta \boldsymbol{r}_{cm}(t)^2 \rangle$*, the blue line the total MSD* $\langle \Delta \boldsymbol{r}(t)^2 \rangle$*, and the red line the segmental MSD in the center-of-mass reference frame* $\langle \Delta \boldsymbol{r}(t)^2 \rangle - \langle \Delta \boldsymbol{r}_{cm}(t)^2 \rangle$*. The red line approaches the asymptotic value* $2R_G^2$ *in the limit* $t \to \infty$*. The dotted line represents the analytical approximation (50), where the integral is* 2.6789*.*

The center-of-mass (47) and segmental (49) MSDs are displayed in Fig. 6. As for the Rouse model, for $t/\tau_Z \ll 1$, the total MSD is dominated by the internal dynamics, i.e., the second term in Eq. (47). Here, the full numerical result and the analytical approximation closely agree with each other, and the MSD increase as $t^{2/3}$. For $t/\tau_R \gg 1$, the MSD is given by the center-of-mass displacement. The deviation between the analytical and numerical solutions in the regime $5 \times 10^{-3} < t/\tau_Z < 10^{-1}$ shows that the $t^{2/3}$ behavior can only be observed for very long polymers and hence a large number of normal modes.

Fluorescence correlation spectroscopy (FCS), a single-molecule technique, provides detailed information on dynamical properties of individual macromolecules [25,30,43]. In Ref. [25], the diffusion and segmental dynamics has been studied of single-end fluorescently labeled dsDNA fragments in the length range from 10^2 to 2×10^4 base pairs (bp), corresponding to $L/l_p \approx 0.7 - 140$. From the correlation function, the mean square displacement of the labeled end is easily extracted. Figure 7 shows the experimental results together with the predictions of the non-draining Gaussian semiflexible polymer model [26, 27, 30, 44]. The theoretical curves are in excellent agreement with the experimental data. To achieve the agreement, an adjustment of the time scale of the model with the diagonal Oseen tensor (41) is necessary. We replaced Ω_{nn} by $\Lambda\Omega_{nn}$ and Ω_{00} by $\Lambda_D\Omega_{00}$, with $\Lambda = 0.6$ and $\Lambda_D = 0.9$ [25]. As shown in Ref. [45], agreement without fit parameter can be achieved, when the full tensor Ω_{nm} is taken into account. Moreover, Brownian dynamics simulations underline the success of the semiflexible polymer model [45, 46]. For short DNA fragments, the polymer dynamics is dominated by the center-

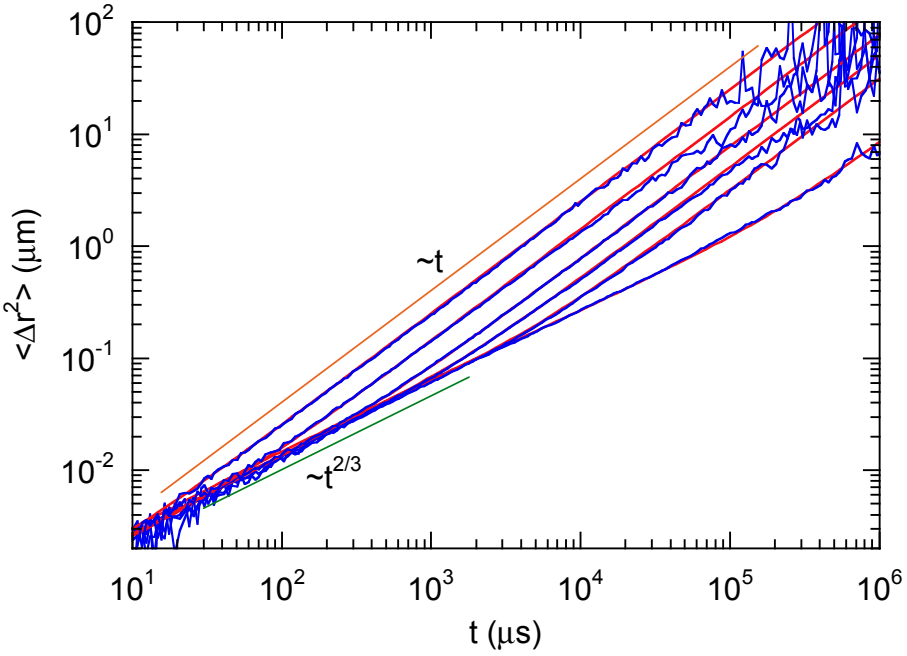


Fig. 7: *Mean square displacements of an end segment of dsDNA molecules of lengths* $L/\mathrm{bp} = 100,\ 200,\ 500,\ 1000,\ 2000,\ 2000$ *(top to bottom). The blue lines are experimental results from Ref. [25], and the red lines are calculated by the semiflexible Gaussian polymer model.*

of-mass motion of the whole molecule in the considered time window. For longer molecules, also the internal dynamics (50) is experimentally accessible. Here, the slope of the MSD is close to $t^{2/3}$ as predicted by theory. The experimental and simulation studies demonstrate that DNA molecules can be considered as semiflexible polymers with their dynamics controlled by hydrodynamic interactions.

The success of the non-draining Gaussian semiflexible polymer model is not a priori evident—the calculations involve a number of approximations, such as the preaveraging approximation. Brownian dynamics simulations reproduce the free-draining behavior less accurately, which suggests that the long-range hydrodynamic interactions contribute to the success of the mean-field model [45].

3.3 Excluded-Volume Interactions

So far, polymers in a Θ solvent have been discussed, i.e., polymers, where the excluded-volume interactions between the various monomers are screened. Experimentally, it has been show that this is an adequate description of polymers in melts (see contribution E 3). In dilute solution under so-called good solvent conditions, excluded-volume interactions lead to a swelling of a polymer coil. The mean square end-to-end distance increases then with the power law $\langle(\boldsymbol{r}_N - \boldsymbol{r}_0)^2\rangle \sim l^2 N^{2\nu}$ with polymer length, where mean field approaches predict $\nu = 0.6$ and more precise renormalization group calculations $\nu \approx 0.588$ [11, 14]. The swelling also affects the dynamics of a polymer. Qualitatively, this effect can be captured by the linearization approximation [14]. Here the distribution function $\Psi(\boldsymbol{r}(s) - \boldsymbol{r}(s')) = F(|\boldsymbol{r}(s) - \boldsymbol{r}(s')|/[|(s-s')/l|^\nu])$

is used, which gives

$$\left\langle \frac{1}{|\boldsymbol{r}(s)-\boldsymbol{r}(s')|} \right\rangle \approx \left(\frac{l}{|s-s'|} \right)^{\nu} \tag{51}$$

and leads to the relaxation times $(L/l \gg 1)$ [14]

$$\tilde{\tau}_n \approx \frac{\eta l^3}{k_B T} \left(\frac{L}{ln} \right)^{3\nu}. \tag{52}$$

With the exponent $\nu = 1/2$ for a Θ solvent, the dependencies (43) are recovered. Renormalization group theory calculations yield the diffusion coefficient $D_Z = 0.2 k_B T/(\sqrt{6}\eta R_G)$ in good solvent, with the radius of gyration $R_G \approx l(L/l)^{\nu}$ [14, 47].

4 Dynamic Structure Factor

The dynamical properties of polymers can be studied by scattering experiments such as dynamic light scattering [48] or neutron scattering (see lecture E 3). Here, the dynamic structure factor

$$S(\boldsymbol{q},t) = \frac{1}{L^2} \int_0^L \int_0^L \langle \exp\left(i\boldsymbol{q}[\boldsymbol{r}(s,t)-\boldsymbol{r}(s',0)]\right)\rangle \, ds ds' \tag{53}$$

is obtained [14]. For the considered polymer model, the average can easily be evaluated, because of the linearity of the model and the Gaussian nature of the underlying stochastic process. Hence, $S(\boldsymbol{q},t)$ becomes

$$S(\boldsymbol{q},t) = \frac{1}{L^2} \int_0^L \int_0^L \exp\left(-\frac{1}{6}\boldsymbol{q}^2 \left\langle [\boldsymbol{r}(s,t)-\boldsymbol{r}(s',0)]^2 \right\rangle \right) ds ds'. \tag{54}$$

With the eigenfunction expansion (19), the exponent can be written as

$$\left\langle [\boldsymbol{r}(s,t)-\boldsymbol{r}(s',0)]^2 \right\rangle = \sum_{n=0}^{\infty} \left[\left\langle \boldsymbol{\chi}_n^2 \right\rangle \left(\varphi_n(s)^2 + \varphi_n(s')^2 \right) - 2 \left\langle \boldsymbol{\chi}_n(t)\boldsymbol{\chi}_n(0) \right\rangle \varphi_n(s)\varphi_n(s') \right]. \tag{55}$$

Equation (54) can be simplified exploiting the equilibrium expression

$$\left\langle [\boldsymbol{r}(s,0)-\boldsymbol{r}(s',0)]^2 \right\rangle = \sigma^2(s,s') = \sum_{n=0}^{\infty} \left\langle \boldsymbol{\chi}_n^2 \right\rangle [\varphi_n(s)-\varphi_n(s')]^2 \tag{56}$$

and the corresponding static structure factor

$$S(\boldsymbol{q},0) = \frac{1}{L^2} \int_0^L \int_0^L \exp\left(-\frac{\boldsymbol{q}^2}{6}\sigma^2(s,s') \right) ds ds', \tag{57}$$

which yields

$$S(\boldsymbol{q},t) = \frac{e^{-\boldsymbol{q}^2 D t}}{L^2} \int_0^L \int_0^L \exp\left(-\frac{\boldsymbol{q}^2}{6} l|s-s'| \right) \times \exp\left(-\frac{\boldsymbol{q}^2 k_B T}{3\pi\eta} \sum_{n=1}^{\infty} \tau_n \varphi_n(s)\varphi_n(s')(1-e^{t/\hat{\tau}_n}) \right) ds ds'. \tag{58}$$

Here, D and $\hat{\tau}_n$ are either the Rouse or Zimm diffusion coefficient and relaxation times, depending on the underlying model. The last term in Eq. (58) describes the internal dynamics of the polymer in the center-of-mass reference frame.

(i) **Small angle regime**: $\boldsymbol{q}^2 lL \ll 1$

For $\boldsymbol{q}^2 lL \ll 1$, the exponents of the terms under the integrals are much smaller than unity and the integrals are well approximated by L^2 (see Fig. 6). Hence,

$$S(\boldsymbol{q},t) = e^{-\boldsymbol{q}^2 Dt}, \tag{59}$$

i.e., the diffusive motion of a polymer coil is measured [14, 20].

(ii) **Large angle regime**: $\boldsymbol{q}^2 lL \gg 1$

To study the internal dynamics of a polymer, the regime $\boldsymbol{q}^2 lL \gg 1$ has to be considered, i.e., the wave length of the scattering vector has to be significantly smaller than the polymer root mean square end-to-end distance. It is sufficient to consider the time regime $t/\hat{\tau}_1 \ll 1$, because $S(\boldsymbol{q},t)$ is very small for $t/\hat{\tau}_1 \gg 1$ [14]. Moreover, the contribution of the translation motion can be neglected, since Dt is much smaller than the contribution from the internal dynamics (see Figs. 3, 6). The integrand in Eq. (58) has a sharp peak at $s = s'$ for the considered large scattering vectors, which yields the expression [20]

$$S(\boldsymbol{q},t) = \frac{2}{L}\int_0^L \exp\left(-\frac{\boldsymbol{q}^2}{6} ls\right) \exp\left(-\frac{\boldsymbol{q}^2 k_B T}{3\pi\eta L}\sum_{n=1}^{\infty} \tau_n \cos\left(\frac{n\pi s}{L}\right)\left(1 - e^{t/\hat{\tau}_n}\right)\right) ds. \tag{60}$$

As for the mean square displacement, the discrete sum can be replaced by an integral.

4.1 Rouse Model

In the integral representation, for the Rouse model follows

$$S(\boldsymbol{q},t) = \frac{12}{q^2 lL}\int_0^{q^2 lL/6} \exp\left(-u - (\Gamma_R(q)t)^{1/2} h\left(u[\Gamma_R(q)t]^{-1/2}\right)\right) du, \tag{61}$$

where ($\gamma = 3\pi\eta$)

$$\Gamma_R(q) = \frac{q^4 l k_B T}{36\pi\eta} \tag{62}$$

and

$$h(u) = \frac{2}{\pi}\int_0^{\infty} \frac{\cos(xu)}{x^2}(1 - e^{-x^2})dx. \tag{63}$$

For $\Gamma_R(q)t \gg 1$, the cosine can be replaced by unity which leads to [14]

$$S(\boldsymbol{q},t) = S(\boldsymbol{q},0)\exp\left(-\frac{2}{\sqrt{\pi}}\left[\Gamma_R(q)t\right]^{1/2}\right). \tag{64}$$

Hence, $S(\boldsymbol{q},t)$ is a universal function of $\Gamma_R t \sim q^4 t$.

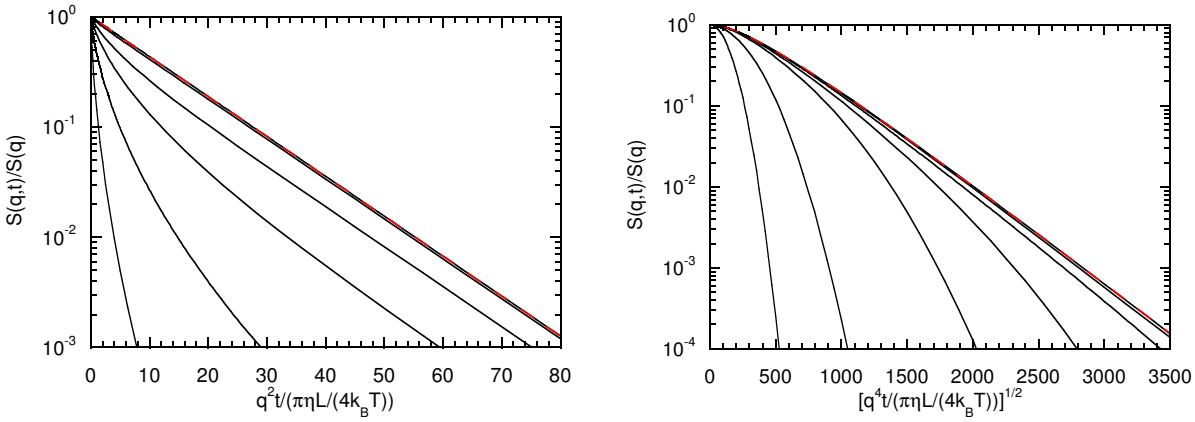


Fig. 8: *Normalized dynamic structure factors of the Rouse model for $L/l = 10^3$ and various q values. The black lines correspond to the full solution of Eq. (58). Left: $S(\boldsymbol{q},t)$ as function of q^2t for $qL = 1$, 100, 200, 300, 500, and 1000 (right to left). Right: $S(\boldsymbol{q},t)$ as function of q^4t for $qL = 50$, 100, 200, 300, 500, 1000, and 3000 (left to right). Note, $q^2lL = 1$ for $qL\sqrt{10^3}$. The red dashed lines indicate the limit (59) for $q^2lL \ll 1$ (left) and (61) for $q^2lL \gg 1$ (right).*

Figure 8 shows dynamic structure factors for $L/l = 10^3$. Hence, for $qL = \sqrt{10^3}$ is $q^2lL = 1$. Then, the conditions $q^2lL \ll 1$ and $q^2lL \gg 1$ translate to $qL \ll 30$ and $qL \gg 30$, respectively. The left figures shows that a universal behavior is obtained for $pL < 30$, where $S(\boldsymbol{q},t)$ is a function of q^2t only. The red dashed line shows the dependence $S(q,t)/S(q,0) = e^{-q^2Dt}$, i.e., the dynamic structure factor decays exponentially in this limit. In the limit $qL \gg 30$, $S(q,t)$ is a universal function of q^4t. As indicated in Fig. 8 (right), the theoretically predicted limit (61) (red dashed line) is assumed for large enough qL-values—in the current case for $qL \gtrsim 10^3$. $S(q,t)$ exhibits the predicted exponential decay (64) for $[q^4t/(\eta L/k_BT)]^{1/2} \gg 1$ only.

4.2 Zimm Model

Similar to the continuum limit (50), the structure factor in the presence of hydrodynamic interactions becomes ($x_n = n(\sqrt{2}t/\tau_Z)^{2/3}$)

$$S(\boldsymbol{q},t) = \frac{12}{q^2lL}\int_0^{q^2lL/6} \exp\left(-u - (\Gamma_Z(q)t)^{2/3}h\left(u[\Gamma_Z(q)t]^{-2/3}\right)\right) du, \tag{65}$$

where

$$\Gamma_Z(q) = \frac{q^3k_BT}{6\pi\eta} \tag{66}$$

and

$$h(u) = \frac{2}{\pi}\int_0^\infty \frac{\cos(xu)}{x^2}(1 - e^{-x^{3/2}/\sqrt{2}})dx. \tag{67}$$

For $\Gamma_Z(q)t \gg 1$ follows [14]

$$S(\boldsymbol{q},t) = S(\boldsymbol{q},0)\exp\left(-1.35\left[\Gamma_Z(q)t\right]^{2/3}\right). \tag{68}$$

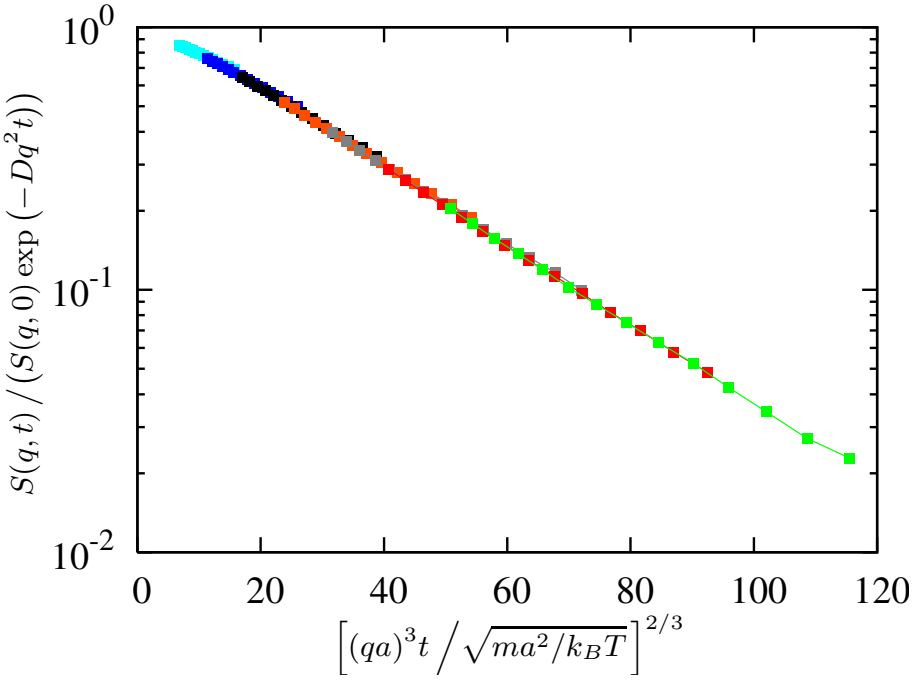


Fig. 9: *Normalized dynamic structure factor of a polymer of length $N = 40$ in dilute solution with excluded volume interactions and for various q values [34].*

Here, $S(\boldsymbol{q}, t)$ is a universal function of $\Gamma_Z t \sim q^3 t$.

There is a clear difference in the dynamic structure factor of the free-draining and non-draining case, which allows us to assess the relevance of hydrodynamic interactions both, experimentally and by simulations.

Our mesoscale hydrodynamic simulations [34] yield the dynamic structure factor displayed in Fig. 9. Independent of the solvent conditions (Θ or good solvent), the Zimm model predicts the scaling relation $S(\boldsymbol{q}, t) = S(\boldsymbol{q}, 0) f(q^3 t)$. Evidently, the simulation results are in agreement with the theoretical prediction, although the function does not decay exponentially at short times. This is similar to $S(\boldsymbol{q}, t)$ of Fig. 8 (right).

5 Conclusions

Naturally, the dynamics of polymers in solution and melt cannot be presented in a comprehensive way in this lecture note. Here, only certain basic aspects have been touched and outlined. There would be much more to be discussed for semiflexible polymers as well as long polymers in melts. In the latter case, the non-crossability of the polymers leads to entanglements and a tube-like confined of a particular polymer by its neighborhood [14] and gives rise to additional power-law regimes in the diffusive dynamics [11, 14].

Appendices

A Oseen Tensor

The Stokes equation of the fluid flow field $\boldsymbol{v}(\boldsymbol{r})$ is given by [19]

$$\eta\Delta\boldsymbol{v} = \nabla p - \boldsymbol{f}, \tag{69}$$

where η is the fluid viscosity, p the pressure, and $\boldsymbol{f}$ an external volume force. The linear equation can be solved by Fourier transformation

$$\boldsymbol{v}_k = \int \boldsymbol{v}(\boldsymbol{r}) e^{i\boldsymbol{k}\boldsymbol{r}} d^3r, \tag{70}$$

which yields

$$\eta\boldsymbol{k}^2\boldsymbol{v}_k + i\boldsymbol{k}p_k = \boldsymbol{f}_k. \tag{71}$$

For an incompressible fluid, where $\nabla\boldsymbol{v} = 0$, follows $\boldsymbol{k}\boldsymbol{v}_k = 0$, i.e., $\boldsymbol{v}_k \perp \boldsymbol{k}$. Multiplication of Eq. (71) with $\boldsymbol{k}$, allows one to eliminate p_k and to obtain the closed expression

$$\boldsymbol{v}_k = \frac{1}{\eta\boldsymbol{k}^2}\left(\mathbf{I} - \frac{\boldsymbol{k}:\boldsymbol{k}}{\boldsymbol{k}^2}\right)\boldsymbol{f}_k,$$

where $\boldsymbol{k}:\boldsymbol{k}$ denotes the tensor product. Fourier transformation gives

$$\boldsymbol{v}(\boldsymbol{r}) = \int \boldsymbol{\Omega}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{f}(\boldsymbol{r}')d^3r', \tag{72}$$

with

$$\boldsymbol{\Omega}(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int \frac{1}{\eta\boldsymbol{k}^2}\left(\mathbf{I} - \frac{\boldsymbol{k}:\boldsymbol{k}}{\boldsymbol{k}^2}\right)e^{-i\boldsymbol{k}\boldsymbol{r}}d^3k,$$

which yields the *Oseen* tensor [14, 19]

$$\boldsymbol{\Omega}(\boldsymbol{r}) = \frac{1}{8\pi\eta r}\left(\mathbf{I} + \frac{\boldsymbol{r}:\boldsymbol{r}}{r^2}\right). \tag{73}$$

References

[1] L. Eichinger, B. Köppel, A. Noegle, M. Schleicher, M. Schliwa, K. Weijer, W. Witke, and P. Janmey, Biophys. J. **70**, 1054 (1996).

[2] P. Janmey, in *Handbook of Biological Physics* (North Holland, Amsterdam, 1995), vol. 1A, chap. 17.

[3] T. L. Hill, *Statistical Mechanics* (MCGraw-Hill, New York, 1956).

[4] H. B. Callen, *Thermodynamics* (Wiley, New York, 1961).

[5] A. Münster, *Statistical Thermodynamics*, vol. I (Springer, Berlin, 1969).

[6] D. Ruelle, *Statistical Mechanics* (W. A. Benjamin, New York, 1969).

[7] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, Oxford, 1987).

[8] P. Reineker, M. Schulz, and B. M. Schulz, *Theoretische Physik V* (Wiley-VCH, Weinheim, 2010).

[9] M. V. Volkenstein, *Configurational Statistics of Polymer Chains* (Interscience, New York, 1963).

[10] L. R. G. Treloar, *The Physics of Rubber Elasticity* (Clarendon Press, Oxford, 1975).

[11] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University, Ithaca, 1979).

[12] J. H. Weiner, *Statistical Mechanics of Elasticity* (Wiley, New York, 1983).

[13] P. J. Flory, *Statistical Mechanics of Polymer Chains* (John Wiley & Sons, New York, 1989).

[14] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Clarendon Press, Oxford, 1986).

[15] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics* (World Scientific, Singapore, 1990).

[16] R. G. Winkler, P. Reineker, and L. Harnau, J. Chem. Phys. **101**, 8119 (1994).

[17] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1989).

[18] H. C. Öttinger, *Stochastic Processes in Polymeric Fluids* (Springer, Berlin, 1996).

[19] J. K. G. Dhont, *An Introduction to Dynamics of Colloids* (Elsevier, Amsterdam, 1996).

[20] L. Harnau, R. G. Winkler, and P. Reineker, J. Chem. Phys. **104**, 6355 (1996).

[21] P.-G. de Gennes, J. Chem. Phys. **55**, 572 (1971).

[22] T. C. B. McLeish, Adv. Phys. **51**, 1379 (2002).

[23] R. G. Winkler, Phys. Rev. Lett. **82**, 1843 (1999).

[24] R. G. Winkler, J. Chem. Phys. **118**, 2919 (2003).

[25] E. P. Petrov, T. Ohrt, R. G. Winkler, and P. Schwille, Phys. Rev. Lett. **97**, 258101 (2006).

[26] R. G. Winkler, J. Chem. Phys. **133**, 164905 (2010).

[27] R. G. Winkler, Phys. Rev. Lett. **97**, 128301 (2006).

[28] R. G. Winkler, Soft Matter **6**, 6183 (2010).

[29] O. Kratky and G. Porod, Recl. Trav. Chim. Pays-Bas **68**, 1106 (1949).

[30] R. G. Winkler, S. Keller, and J. O. Rädler, Phys. Rev. E **73**, 041919 (2006).

[31] P. E. Rouse, J. Chem. Phys. **21**, 1272 (1953).

[32] B. H. Zimm, J. Chem. Phys. **24**, 269 (1956).

[33] M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper, Europhys. Lett. **68**, 106 (2004).

[34] K. Mussawisade, M. Ripoll, R. G. Winkler, and G. Gompper, J. Chem. Phys. **123**, 144905 (2005).

[35] S. R. Quake, H. Babcock, and S. Chu, Nature **388**, 151 (1997).

[36] T. A. King, A. Knox, W. I. Lee, and J. D. G. McAdam, Polymer **14**, 151 (1973).

[37] M. Adam and M. Delsanti, Macromolecules **10**, 1229 (1977).

[38] T. Nose and B. Chu, Macromolecules **12**, 590 (1979).

[39] C. C. Han and A. Z. Akcasu, Macromolecules **14**, 1080 (1981).

[40] W. H. Stockmayer and B. Hammouda, Pure & Appl. Chem. **56**, 1373 (1984).

[41] M. Fixman, J. Chem. Phys. **78**, 1594 (1983).

[42] B. Liu and B. Dünweg, J. Chem. Phys. **118**, 8061 (2003).

[43] E. P. Petrov and P. Schwille, in *Standardization and Quality Assurance in Fluorescence Measurements II*, edited by U. Resch-Genger (Springer Berlin Heidelberg, 2008), vol. 6 of *Springer Series on Fluorescence*, p. 145.

[44] R. G. Winkler, J. Chem. Phys. **127**, 054904 (2007).

[45] M. Hinczewski and R. Netz, EPL **88**, 18001 (2009).

[46] M. Hinczewski, X. Schlagberger, M. Rubinstein, O. Krichevsky, and R. R. Netz, Macromolecules **42**, 860 (2009).

[47] Y. Oono and M. Kohmoto, J. Chem. Phys. **78**, 520 (1983).

[48] B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley, New York, 1976).

C 1 Neutron Sources

A. Ioffe
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Neutron scattering is a very important tool for studies of fundamental properties of condensed matter as well as material research. It has a special stand among other kinds of radiation as light, X-rays or synchrotron radiation, electrons or ions because of electrical neutrality of neutrons, its large magnetic moment and low kinetic energy. Due to these unique properties of neutrons they became irreplaceable for the investigations of static and dynamic properties of condensed matter, magnetic properties and living biological objects.

The performance of neutron scattering instruments, i.e. the precision of carried out experiments, is primarily determined by the recorded intensity of the scattered beam. The latter is proportional to the unit scattering power of a sample σ_s (the scattering cross-section), to its volume V_s and to the incident neutron flux I_0:

$$I_{\det} = \varepsilon_{pr}\varepsilon_{\sec}\varepsilon_{\det}\sigma_s V_s I_0 \qquad (1)$$

where ε_{pr} , ε_{sec} and ε_{det} are efficiencies of primary, secondary spectrometers and detector system, respectively (discussions of these elements of neutron spectrometers will take place in dedicated lectures) (Fig.1). A general tendency in modern science is to investigate smaller samples (such as nanostructures, biological objects, etc.) and weaker effects, so that the unit scattering power of a sample σ_s and the sample volume V_s are very small. Indeed, the flux at the neutron scattering instrument becomes an ultimate parameter that defines the quality of the experiment.

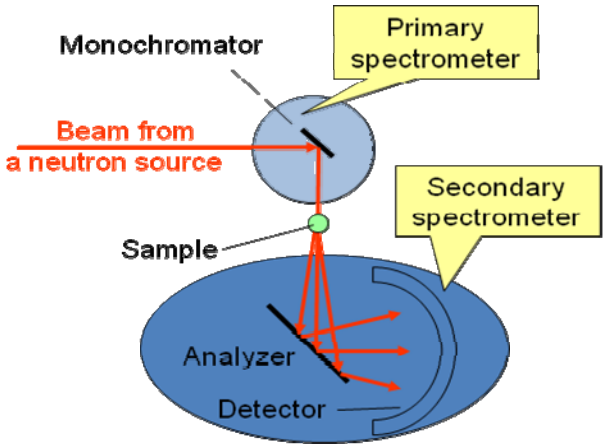


Fig. 1: *Layout of a neutron spectrometer.*

Usually, under the neutron source one understands a nuclear installation emitting neutrons. However, from the point of view of neutron scattering the neutron source should be considered more generally, including also a spectrum transformer, tailoring the neutron spectrum according to the parameters of the neutron scattering instruments and the neutron

transport system that delivers neutrons to the instrument sites (Fig. 2). In this lecture we will discuss all these three components in more details.

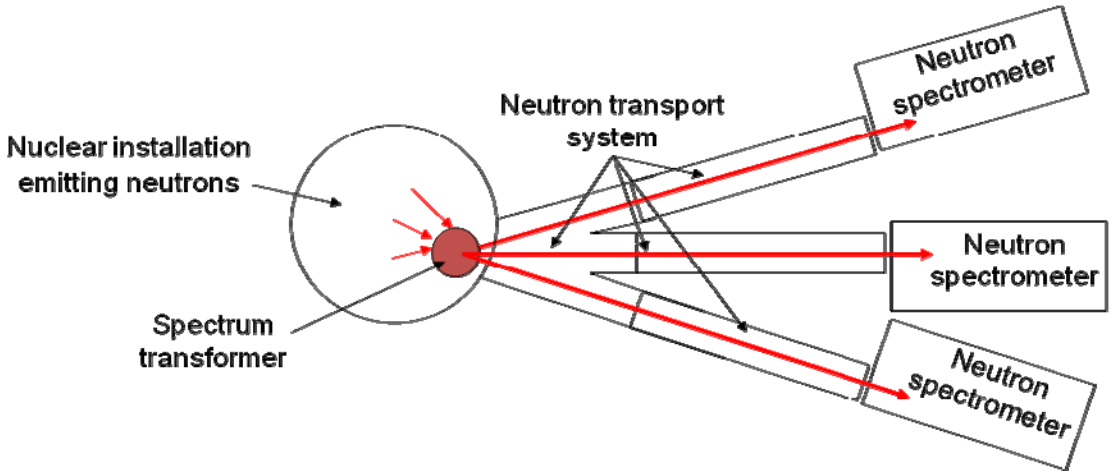


Fig. 2: *Layout of a neutron scattering facility.*

2 Nuclear Reactions

In nature neutrons are strongly bound in the atomic nuclei. Therefore, despite the fact that neutrons constitute about a half of each atom so that nature for a half is comprised of neutrons, it is rather difficult to set them free. Therefore, the only way to free neutrons from the nuclear confinement is to break a nucleus apart by means of a nuclear reaction. Further we will consider two types of such reactions – fission and spallation nuclear reactions – that are used in modern continuous and pulsed neutron sources, respectively.

2.1 Nuclear fission reaction

Namely by this way, bombarding the beryllium nuclei with of α-particles obtained from decay of natural polonium, Chadwick has produced the first free neutrons in 1932:

$$^{9}_{4}\mathrm{Be} + ^{4}_{2}\alpha \rightarrow ^{12}_{6}\mathrm{C} + ^{1}_{0}\mathrm{n}$$

However, the neutron flux available from such sources was far away from being useful for the condensed matter investigation. The breakthrough happened in the 40^{ies}, when nuclear reactors using the nuclear fission reactions have been constructed. Although these reactors have been primarily developed for purposes of the nuclear weapon industry, a by-product of their operation - an enormous for that time neutron flux, about 10^7 neutrons per square centimetre per second (n/cm^2 s) at the CP-1 reactor in USA - was immediately used for first neutron scattering experiments. Developments in technology of fission reactors during the next 30 years resulted in a tremendous, by 8 orders of magnitude increase of the neutron flux of nuclear research reactors that approached 10^{15} n/cm^2 s for the high-flux reactor of the Institut-Laue-Langevin (Grenoble, France) in 1972.

These reactors are using the fission of the uranium isotope ^{235}U. Following the capture of a slow neutron, this nucleus is deformed and is split into two fragments, simultaneously releasing 2 or 3 (on average 2.5) "prompt" neutrons with the average energy of about $E_T \approx 2$ MeV (Fig. 1):

$$^{235}\text{U} + \text{neutron} \rightarrow \text{fission fragments} + 2.52 \text{ neutrons} + 180 \text{ MeV}. \qquad (2)$$

Each of these practically instantly (within 10ns) emitted neutrons can cause the fission of another 2-3 nuclei, so that each of them will also emit 2 to 3 neutrons, and so on (see Fig. 3). This process is called the chain reaction, where the amount of fissile material needed to sustain the chain reaction is called critical mass M_c. If the mass M of fissile material is more than critical, $M>M_c$, the number of neutrons will increase exponentially and the reaction will become uncontrollable very quickly, leading to a huge energy release. If the mass of fissile material is less than critical, $M<M_c$, it will be impossible to sustain a chain reaction: the number of neutrons will decrease over time. Thus this neutron producing reaction is unstable and will not provide a stable neutron flux.

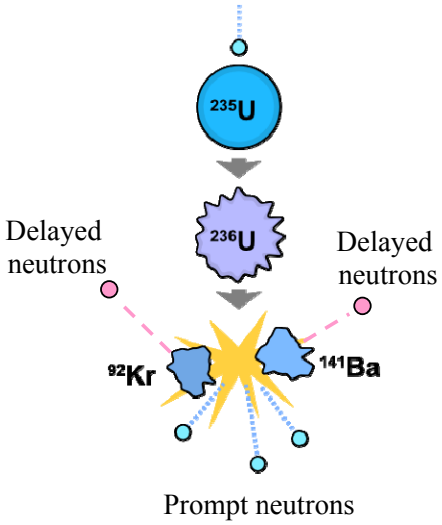


Fig. 3: *Schematic representation of the fission process of U-235.*

How to obtain a stable neutron flux? Fortunately, there is another additional mechanism that saves the situation - the fission fragments are also rich in neutrons and emit neutrons as a part of their radioactive decay, which can also contribute to the fission of any U-235 nucleus they strike. These so-called "delayed" neutrons, though they make only about 0.64% (!) of the total amount, are extremely important because they are emitted with the average time delay of the order of seconds and thanks to them the chain reaction can be controlled. Practically one runs a reactor sub-critically as far as only prompt neutrons are concerned, i.e. neutron multiplication is suppressed, so that the chain reaction vanishes. The delayed neutrons come a moment later but just in time to sustain the chain reaction when it is going to die out, thus allowing to reach criticality (see Fig. 4). More precisely, the neutrons in the reactor are moderated to decrease their energy and to increase their absorption by control roads that are made of a neutron absorbing material (usually containing boron). When inserted in the reactor

core, these roads will reduce the number of slow neutrons to the amount just as necessary for the self–sustaining chain reaction and may be adjusted, so that the reaction remains critical only with the inclusion of the delayed neutrons. Thus, a simple and reliable mechanical control system can be used for the control of the chain reaction in the nuclear reactor.

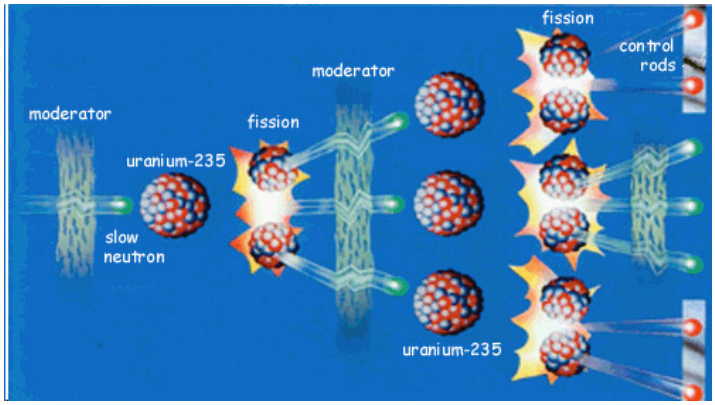


Fig. 4: *Controlled chain reaction in the nuclear reactor. Control rods reduce the number of slow neutrons to the amount just as necessary for the self–sustaining chain reaction. By the proper adjustment of the control rods' position, the reaction may remains critical only with the inclusion of the delayed by a few seconds neutrons.*

A change in the reactor power results in changes in temperatures of its fuel. For example, when the power rises, the temperature of the uranium fuel will rise as well. However, the higher the temperature, the higher the ability of U-238 to absorb neutrons. Indeed, a mass of a fissile material that is exactly critical at room temperature becomes sub-critical if it is warmed and the chain reaction dies out without any external interaction. This so-called negative coefficient of reactivity is an inherent safety factor of nuclear reactors.

2.2 Spallation reaction

The fission is not the only nuclear reaction allowing us to obtain free neutrons. Another kind of nuclear reactions that can be used for neutron production is the spallation reaction (Fig.5), where extremely high energy particles (e.g. protons) hit the target made of a neutron-rich material, "breaking" a heavy nucleus into highly excited fragments. In contrast to the fission reaction, the de Broglie wavelength

$$\lambda = \sqrt{h^2/2mE} \tag{3}$$

of the bombarding particles with energy E is shorter than the size of the nuclei, and collisions can take place with individual nuclides in the nucleus rather than with the nucleus as a whole (here h is the Plank constant, m is the neutron mass) . Indeed, a large amount of energy is transferred to the nuclides, which in turn can hit other nuclides in the same nucleus. As the result of this so-called intranuclear cascade, energy is more or less evenly distributed over the nucleus, bringing it to a highly excited state, so that the excited nucleus will "evaporate"

neutrons and a smaller amount of protons. However, some energetic particles can escape from the nucleus and either hit another one (internuclear cascade) or just escape from the target. The energy of these neutrons is extended up to the energy of the incident particles (i.e. up to 1 GeV).

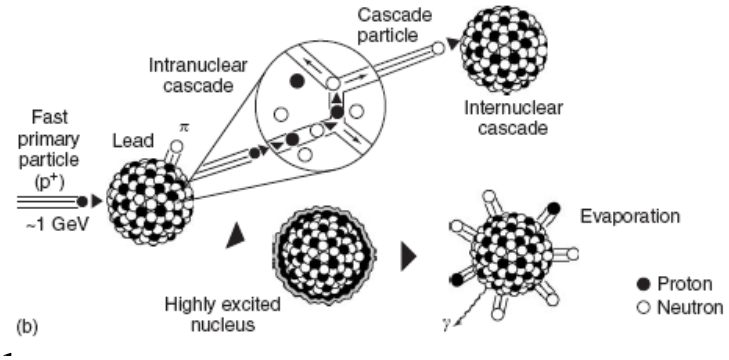


Fig. 5: *Schematic representation of the spallation process.*

The spallation process is very short and ends within less than 10^{-15} s after the nucleus is hit. Thus, the time distribution of the spallation neutrons is entirely determined by the time distribution of the driving particle pulse, generally provided by a linear accelerator. This pulse that can be made either rather long, about 5 ms (called the long pulse spallation source (LPSS)), or rather short, about 10 μs (called the short pulse spallation source (SPSS)), by compressing charged particles in a compressor ring.

The most intense up-to-day spallation neutron source ISIS at Chilton (Great Britain) provides instantaneous thermal neutron fluxes over 10^{16} n/cm^2 s with short pulse lengths of ~50 μs. Next generation of pulsed neutron sources – SNS in USA and JHP in Japan - with fluxes more than 10^{17} n/cm^2 s have started their operation and gradually approaching their projected parameters. The European project of 5MW spallation source ESS with flux more than 10^{17} n/cm^{2}s is expected to take off within the next years.

Comparing possible nuclear reactions that can be used for neutron production (see Table 1), one should pay attention not only to their efficiency. The heat deposition that accompanies the neutron production results in the cooling problem, which is the real limiting factor for all kinds of neutron sources. From this point of view, fusion is the most attractive process, although it is still a technique of a far future; in the same time, spallation is more attractive than fission.

Table 1. *Neutron yields and deposited heat for selected neutron-producing reactions.*

Reaction	Energy/event	Yield (neutron/event)	Deposited heat (MeV/neutron)
(T,d) fusion		~1 neutron/fusion	3
^{235}U fission		~1 neutron/fission	200
Pb spallation	1 GeV	~20 neutron/proton	23
^{238}U spallation	1 GeV	~40 neutron/proton	50

3 Neutron spectrum and spectrum transformation

To be useful for condensed matter investigations neutrons wavelength λ should be about few Angstroms, that corresponds to the meV energy range. However, the energy spectrum of neutrons produced by fission or spallation nuclear reactions is in the range of 1 MeV, i.e. the spectrum transformation aiming an energy shift of several orders of magnitude is required. Let us express the neutron energy in the terms of the mean temperature T of the neutron ensemble as $E=k_BT$, where k_B is the conversion coefficient $k_B = 8.617\times10^{-5}$ eV/K (the Boltzmann constant): then one can say that required energy shift can be achieved by cooling the neutrons down to a much lower temperature. For this purpose neutrons should be brought into the thermal equilibrium with a cold body (moderator): because of multiple inelastic collisions, like billiard balls, with the light atoms (the mass A) of the moderator neutrons are slowing down with the mean logarithmic reduction of neutron energy $\xi = \ln\frac{E_T}{E} \approx \frac{2}{A+1}$ per collision till $E = E_M = k_BT_M$ (T_M - the moderator temperature), thus achieving thermal equilibrium $T \approx T_M$ with the moderator within 10^{-6} s after $n = \frac{1}{\xi}\left(\ln E_T - \ln E_M\right)$ collisions. This is the so-called thermalization process, which as one can see from the above-mentioned formula most effective for the smallest A (it takes 16 or 29 collisions to bring a neutron from the fission energy of $E_T = 2$MeV to an $E_M = 1$eV for H_2O and D_2O, respectively). From other hand, depending on the type of neutron source the moderator should provide either the highest possible flux in the largest possible volume (for continuous neutron sources) or in the shortest possible time (for pulsed neutron sources). This can be achieved by using water (A=1) or heavy water moderators (A=2). The neutron energy spectrum is given by the Maxwellian distribution

$$\Phi(E) = \frac{2\sqrt{E}}{\sqrt{\pi k_B{}^3 T_M{}^3}} \exp\left\{-\frac{E}{k_B T_M}\right\} \tag{4}$$

Practically, moderators are big (light or heavy) water volumes (also serving as a biological shielding) surrounding the reactor core or the spallation target and are generally kept at the room temperature of $T_M \approx 300$ K. Because of this reason the corresponding neutrons are called thermal neutrons, with a maximum peak flux around $\lambda \approx 1$ Å (see Fig. 6).

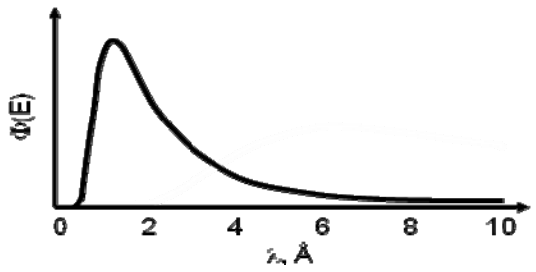


Fig. 6: *The energy distribution for neutrons produced by a neutron source for the moderator temperature T_M=300K.*

Tables 2 and 3 contain some neutron properties and useful relations between different parameters of neutrons.

Table 2. *Neutron properties*

Mass	$m = 1.675 \cdot 10^{-27}$ kg
Electrical charge	$q = 0$
Magnetic dipole moment	$\mu_n = -1.913\ \mu_B$ (μ_B – nuclear magneton)
Life time	$t_{1/2} = 820$ s

Table 3. *Some useful relations*.

$\lambda(\text{Å}) = \frac{h}{mv} = \frac{3956}{v(m/s)} = \frac{0.286}{\sqrt{E(eV)}}$	$v(m/s) = \frac{h}{m\lambda} = \frac{3956}{\lambda(\text{Å})}$

4 Nuclear reactor and spallation source

Now we can consider the construction of a neutron source. In all cases its heart is a core where the nuclear fission reaction takes plays. In the case of the nuclear reactor a set of uranium ^{238}U fuel elements (or a complex single fuel element) is enriched by the isotope ^{235}U. One distinguishes between high-enriched (~95%) and low-enriched (~20%) uranium: because the amount of fission material necessary to support the chain reaction is predetermined (the critical mass), the enrichment of the used uranium fuel actually defines the volume of the core and the neutron flux density. Compact cores made of high-enriched uranium at the reactors of the Institut-Laue-Langevin (Grenoble, France) and the Maier-Leibnitz Neutron Source (FRM-2) (Garching, Germany) provide the highest flux density and therefore, the highest neutron source luminosity achievable up to this day. The reactor core is surrounded by (heavy) water (T=300 K) that plays the role of the moderator of high-energy fission neutrons (Fig. 7). Obviously, the full thermalization of these neutrons requires some time necessary for a few collisions with hydrogen (or deuterium) atoms, so that the density of thermal neutrons increases with the distance r_0 from the core. On the other hand, the neutron absorption is inverse proportional to the neutron velocity, so that the flux of already thermalized neutrons decreases with r_0. As the result of these two competing processes, the thermal neutron flux density achieves its maximum at a certain distance of r_0 = 10-15 cm from the core (Fig. 8). Obviously, to extract thermal neutrons from the reactor, the entrance of a neutron beam tube should be placed exactly in this position. Aiming the decrease in undesirable background of fast (i.e. still not thermalized) neutrons and γ-rays from the core, one should avoid the direct view of the core through the neutron beam tube. All together, it leads to the conclusion that the optimal arrangement of beam tubes is tangential to the reactor core (see Fig. 7).

In case of the spallation source, the role of the reactor core plays a target made of heavy metal as Bi, Pb or Hg. The proton beam is obtained from negatively charged hydrogen ions produced by powerful ion sources (Fig. 9): each ion consists of a proton orbited by two electrons. These ions are accelerated in a linear accelerator (Linac) by a number of subsequent radio-frequency cavities with strong electro-magnetic fields, thus achieving kinetic energies

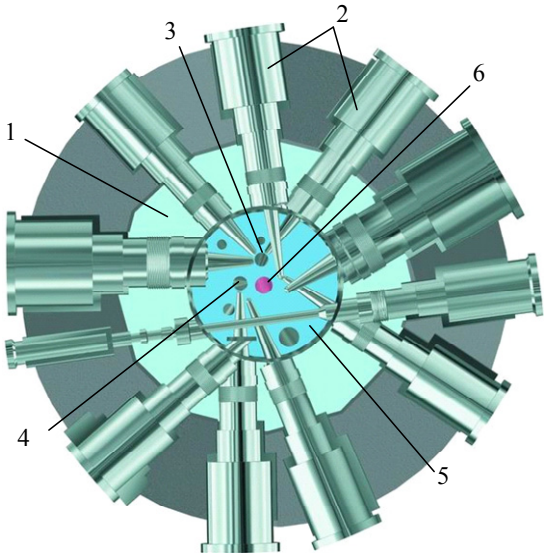


Fig. 7: *Horizontal section through the reactor pool of the FRM-2 reactor in Garching, Germany. The reactor tank with internal diameter approx. 5m is filled with light water (1). In the centre of the arrangement the reactor core is situated. The experimental installations as horizontal beam tubes (2), a cold (3) and a hot (4) neutron source are arranged in the heavy water tank (5) around the fuel element (6).*

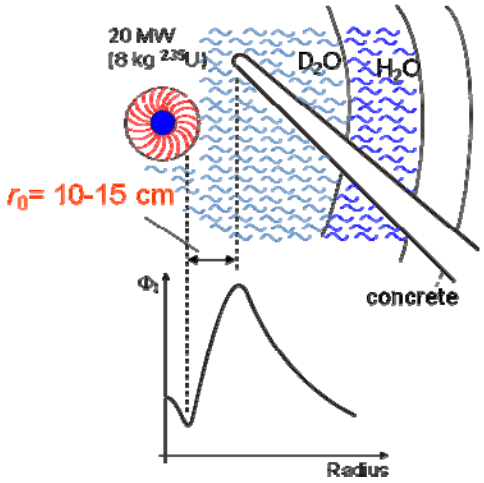


Fig. 8: *The density of thermal neutrons vs. the distance r_0 from the core and the tangential arrangement of beam tubes in the reactor core*

in the GeV range (i.e. about 90% of the speed of light). When these hydrogen ions leave the Linac, they are stripped off all their electrons by passing through a thin carbon sieve, so that the negative hydrogen ions become protons. Now, depending on the design of the spallation source, LPSS or SPSS (see Chapter 2) the protons are either sent to the target directly or

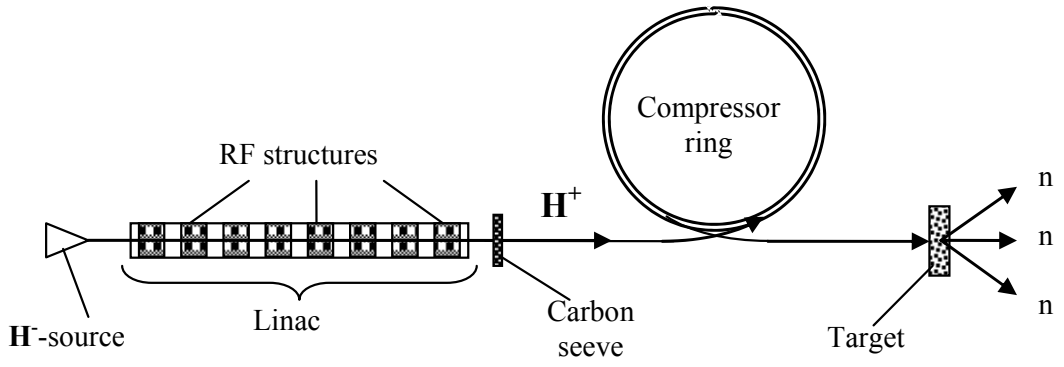


Fig. 9: *Layout of the neutron spallation source.*

through a compressor ring, respectively. The latter collects the protons from a large number of successive bunches from the Linac into a single very high-intensity proton pulse. It is achieved by an assembly of magnets that send each accelerated proton bunch into a circular orbit of such a large diameter (~50–100 m), so that the travelling time is equal to the time interval between the bunches. Indeed, the next bunch of protons arrives exactly when the previous has made a full turn and both of them are sent around again. When about 1000 bunches are piled up by such a procedure, sufficient intensity is accumulated and the full proton pulse with a pulse length of about 1 μs is sent to the target. The target is normally a liquid metal (mercury or a lead–bismuth eutectic mixture), placed in special materials to consume the beam power of a few megawatts. The proton pulse repetition rate on the target should be about 10–100 Hz to achieve an optimal use of neutrons in time-of-flight scattering experiments.

Thus, there are two kinds of neutron sources and certainly the question arises, which of them is better answering future trends. These trends, as discussed in the Ch.1 require a significant increase in the luminosity of neutron sources in order to improve the counting statistics of neutron scattering experiments. However, the evolution of nuclear reactors that was very impressive some decades ago, shows no progress since 1972, when the high-flux reactor at the ILL, Grenoble became operative. The reason for this is clearly the technical difficulty of removing the heat from the reactor core.

Let us make some rough estimates. As it was mentioned in Table 1, the deposited heat amounts to 200 Mev/fission with the yield of 1 neutron from 2.5 to be extracted for neutron scattering experiments. Using the relation 1 eV = $1.6 \cdot 10^{-19}$ J, we obtain the source strength (i.e. the number of neutrons emitted per second) $Q = 3 \cdot 10^{16}$ n/s per MW of the reactor power to be removed. However, this is a kind of a "point neutron source" that immersed into the moderator to slow neutrons down till thermal energies (see Ch. 2). Indeed, all neutrons emitted by the point source will be spread over the moderator surface of about 2000 cm^2 (r =10-15 cm), so that the thermal neutron flux will amount to 0.0005 of Q, i.e. about $1.5 \cdot 10^{13}$ $n/s \cdot cm^2$ per MW of reactor power. Thus, for the 57 MW reactor at the ILL, one may expect the thermal flux of $1 \cdot 10^{15}$ $n/s \cdot cm^2$ to be compared with the actual value of $2 \cdot 10^{15}$ $n/s \cdot cm^2$.

Thus, a further increase in the thermal neutron flux from nuclear reactors will require a significant increase in their power. However, such an increase will also require a very sophisticated reactor cooling and result in even stronger radiation damage of the reactor vessel components (beam tube noses, cold source, etc.). Experience gained at the ILL reactor shows that their service time is seven years. Already now new reactors are being designed in a way allowing for a regular exchange of the beam tube noses. Tenfold increase of the reactor power will result in a rather unpractical service time of these elements. Another problem is the worldwide concern about a potential risk associated with nuclear fission installations. On the other hand, pulsed sources are inherently safer because of the absence of any critical configuration that is potentially explosive. The deposited heat is 10 times less with the simultaneous significantly large neutron output (see Table 1) allows for a high peak flux about 50 times higher than the one for the ILL reactor (Table 2). Losses in the average thermal neutron flux will be compensated by the opportunities offered for neutron scattering instrumentation by the time-structured neutron beams, when the instrument performance depends on the peak flux in the pulse rather than on the time-averaged flux.

Therefore, it is not surprising that all new sources under construction, SNS in USA and JHP in Japan as well as planned new European neutron source ESS, are spallation neutron sources. However, SNS and JHP are 1–2 MW spallation sources designed to create rather short neutron pulses of about 100μs and further increase in their power level is rather problematic due to possible target problems. In contrast to this, ESS is planned as 5MW spallation source because it will create neutron pulses of a few milliseconds duration. It was demonstrated that such long pulse provides significant advantages for certain categories of neutron scattering instruments. Although target problems for LPSS also become increasingly severe with the increase of power, nevertheless it seems realistic to approach the ultimate limit of 20MW (i.e. 20 mA proton current at 1 GeV).

5 Cold, thermal and hot neutrons

As it was shown in the previous Chapter, the energies of neutrons produced by neutron sources cover many orders of magnitude. Depending on their energy E, neutrons are classified by commonly used names (see Table 4). Neutrons with $E < 1$ keV are called slow neutrons; in turn they are classified in 6 groups, but the most relevant for purposes of neutron scattering

Table 4. *Classification of neutrons according to their energy (wavelength).*

Ultra cold	E <0.5 μeV	λ > 400 Å	T < 6mK
Very cold	E=0.5μeV-0.05 meV	λ = (40-400) Å	T = 6mK-0.6°K
Cold	E=(0.05-5) meV	λ = (4-40) Å	T = (0.6-60)°K
Thermal	E=(5-100) meV	λ = (0.9-4) Å	T = (60-1000)°K
Hot	E=100 meV -1eV	λ = (0.3-0.9) Å	T = (1000-10000)°K

are hot, thermal and cold neutrons - energy ranges corresponding to these groups are also presented in the water moderator and is about $\lambda \approx 1$ Å (see Ch. 3). A one can see from Fig. 10, most of the neutrons are concentrated around this wavelength in the range of (0.8÷2) Å. These neutron wavelengths perfectly match the interatomic distances in solids and therefore, are extensively used for the studies of structure and dynamics of crystalline.

However, it also means that the amount of hot or cold neutrons in the thermal neutron spectrum is very small, so that any scattering experiment which requires hot or cold neutrons will suffer from enormous flux losses. To enhance the hot or cold neutron flux one has to transform the thermal neutron spectrum shifting it towards high or low energies: in other words by heating or cooling the thermal neutron spectrum.

To achieve a significant gain factor by such spectrum transformations, the moderator temperature should be as high as 2000K and as low as 20K, respectively. Obviously, it is unrealistic to heat or to cool the whole water in the reactor vessel - tens of cubic meters – to such temperatures. The trick that is used to solve this problem is to insert other small local moderators inside the water and to set their temperatures accordingly. These devices are called hot and cold sources.

The hot source is usually made of a graphite block heated up to T=2400 K, when the cold sources is usually a vessel filled with liquid H_2 or D_2 or their mixture cooled down to 20K. Hot and cold neutron spectra are shown in Fig. 10. Each of them allows for a significant, up to 20 times gain in the corresponding neutron flux. By choosing the adequate neutron spectrum scattering experiments can be optimally tailored to particular experimental requirements.

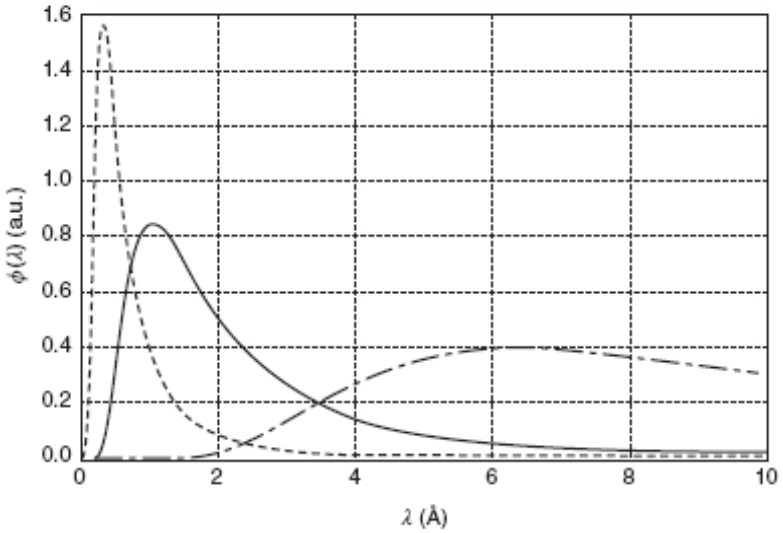


Fig. 10: *Neutrons wavelength distribution from cold (dot dashed line, T=50 K), thermal (solid line, T=300K) and hot (dashed line, T=1000 K) moderators***.**

6 Neutron beam transport

However, it is not enough to produce neutrons in the moderator – they still have to be transported to a neutron scattering instrument. As it was already mentioned in Ch. 4, neutrons are extracted from the moderator by neutron beam tubes, inserted in a heavy biological shielding surrounding the reactor tank and necessary because neutrons are isotropically emitted from the moderator.

The angular acceptance of a neutron beam tube is defined by its diameter (~10 cm) and length (~5m). Thus, the beam divergence of a beam tube is about ~1°, so that the neutron flux available at its output is drastically reduced by about six orders of magnitude, in comparison to the core flux.

This situation can be significantly improved by using neutron optical devices called neutron guides. The principle of their operational is rather similar to the one of light guides, where the light propagating in an optically dense media (i.e. with the refraction index n >1) is totally reflected from the glass air-interface due to the effect of total external reflection (the refraction index of air is equal to unity). In contrast to light, the refraction index of glass for neutrons is n <1, so that the effect of the total external reflection will take place on the air-glass interface. However, in case of neutrons this phenomenon takes place only for incident angles, i.e. less than the critical angle θ_c , which is given by

$$\theta_c = \lambda\sqrt{\frac{2\rho b_{coh}}{\pi}} \tag{5}$$

where ρ and b_{coh} are the density and the coherent scattering length of the wall material, respectively. To increase θ_c the Ni coating with the critical angle $0.1° \cdot \lambda$ is used. Moreover, the wall of the neutron guide can be coated with so-called supermirrors, with the critical angle up to three times as much as the nickel's one. Indeed, the neutron guide is made as a hollow glass tube, Ni or supermirror coated from the inside (Fig. 11). Because the intensity at the neutron guide output is proportional to θ_c^2, they provide an order of magnitude flux increase as compared to a beam tube.

Fig. 11: To principle of the operation of light guides (a) and neutron guides (b)

Moreover, neutron guides can be bent or shaped. Bent neutron guides allow to avoid direct sight-of-view of the reactor core, drastically reducing γ- and neutron background at the instrument position. The parabolic or elliptic shaping of neutron guides opens exciting

possibilities for the concentrating (focusing) of neutrons on a sample, thus providing additional increase in intensity at the position of neutron scattering instruments.

References

[1] G.R. Bauer (1993) Neutron sources. In: A.Furrer (ed.) *Neutron Scattering*, pp. 331–357. PSI-Proceedings No. 93-01, ISSN 1019-6447, Paul Scherrer Institute, Villigen.
[2] J.M. Carpenter and W.B. Yelon (1986) Neutron sources. In: K. Sköld and D.L. Price (eds.) *Methods of Experimental Physics*, vol. 23A, pp. 99–196. London: Academic Press.
[3] K. Clausen (2001) *Neutron sources*. Office for Official Publications of the European Communities, Luxembourg, ISBN 92-894-0037-4.
[4] Utsuro, Masahiko and Ignatovich, Vladimir, *Handbook of Neutron Optics*. ISBN-13: 978-3-527-40885-6 - Wiley-VCH, Berlin.

C2 Synchrotron Radiation Sources

Wolfgang Eberhardt
DESY-Center for Free Electron Laser Science (CFEL),
Notkestr. 85, 22607 Hamburg
and
TU-Berlin, Institut für Optik und atomare Physik,
Straße des 17. Juni 135, 10623 Berlin

Contents

1. Introduction to Synchrotron Radiation

Synchrotron radiation was discovered in 1947 at an electron accelerator – a synchrotron - operated in the laboratories of General Electric in the USA, shortly after its properties had been described by Schwinger (see ref.1). Generally, synchrotron radiation is the characteristic emission of relativistic charged particles accelerated by a magnetic field. Accordingly it not only is observed in accelerators but it also is part of the radiation spectrum observed in astronomy. In space there are a lot of highly accelerated charged particles and when these particles are deflected (accelerated) by powerful magnetic fields, of pulsars for example, they emit electromagnetic radiation, which exhibits the characteristics of synchrotron radiation (see chapter 2).

The synchrotron radiation emitted by an electron accelerator has a brilliance, that is the number of photons emitted per unit area of source size into a specific solid angle within a given wavelength band, which is by 10-12 orders of magnitude superior to the brilliance of laboratory X-ray sources. The brilliance distinguishes a laser from a light bulb. Obviously the laser is superior, since more photons are offered within a given wavelength range, spot size, and beam divergence. The brilliance of a light source is a unique quality factor, which cannot be improved by any optics. Tighter focusing gives a smaller spot size, but the divergence of the beam increases proportionally. This superior brilliance enables scientific investigations that are simply not possible to perform with laboratory sources. This is the origin of the phenomenal success and the unique contributions of accelerator based light sources to scientific discovery, which covers many fields from physics, chemistry, and materials sciences to engineering, biology and life sciences. Even historical sciences and archaeometry have recently benefitted from the unique, non-destructible characterization techniques offered by synchrotron radiation.

The brilliance achieved at various synchrotron radiation lights sources and the historical development is shown in Fig. 1

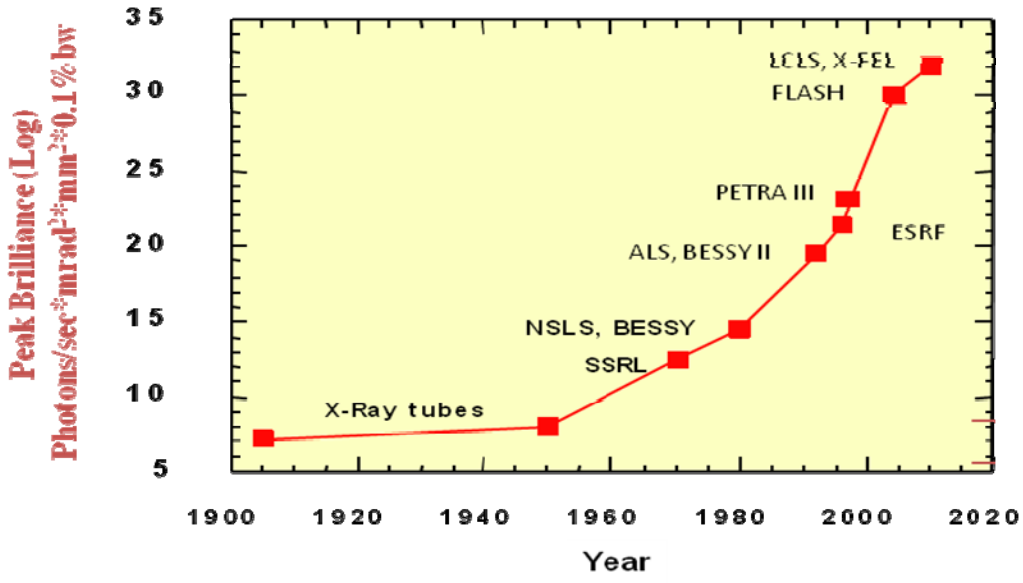


Fig. 1: *Peak brilliance of laboratory X-ray sources, 1st (SSRL), 2nd (NSLS, BESSY), and 3rd generation (ALS, BESSY II, ESRF, PETRA III) synchrotron radiation sources, and Free Electron Lasers (FLASH, LCLS, X-FEL)*

Initially synchrotron radiation experiments were carried out on particle accelerators constructed for high energy particle physics experiments in a parasitic mode. Here the synchrotron radiation is (an unwanted and un-welcome) by-product, which limits the ultimate energy that can be reached by a particular accelerator. Nevertheless specific laboratories were built to use these so called first generation sources. As the success of these experiments became visible and the demand for this quality of radiation increased accelerators were designed exclusively for the use as light sources. These were the 2nd generation sources. Accordingly their parameters were tuned to optimal performance as a light source where for example the electron beam should be quite small all around the circumference of the ring, as opposed to in an accelerator for colliding beam experiments, where the interaction is maximized at a few interaction points around the ring. The next generation of sources incorporated specific magnet structures as radiation sources, so called wigglers and undulators (see section 2), which were introduced into straight sections of the accelerators, further enhancing the light source properties, especially the brilliance, by orders of magnitude. The ultimate gain in brilliance now results in the Free Electron Laser sources, where only a few facilities are operational around the world.
The history of the development of synchrotron radiation and the initial sources around the world has been very nicely described by Arthur L. Robinson in the X-Ray data booklet (1) which is also available in the internet.

2. Properties of synchrotron radiation

Any electron that is accelerated emits electromagnetic radiation, which in the most general term is called 'Bremsstrahlung'. This is the continuous broad range background which is observed together with the characteristic elemental line spectrum of any laboratory X-ray source. The lines are due to atomic transitions observed in the decay of core electron vacancies excited by the electron beam hitting the X-ray anode and thus their energy is characteristic of the anode material.

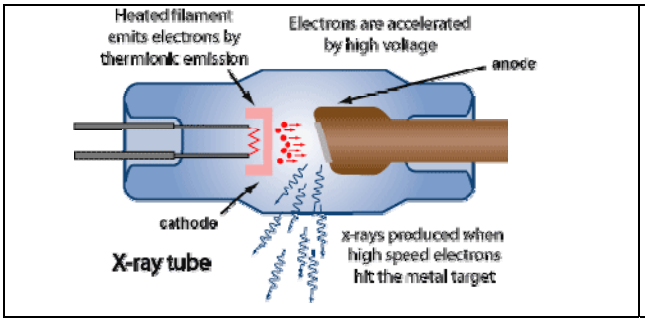


Fig. 2: *Schematic layout of an X-ray tube*

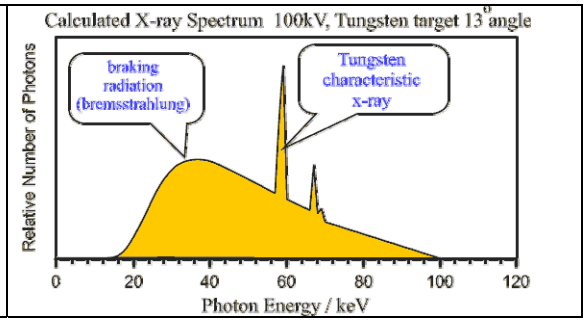


Fig. 3: *Calculated X-ray spectrum of a W target under 100 keV electron beam radiation*

Source: Australian Radiation Protection and Nuclear Safety Agency (ARPANSA)

Compared to an such X-ray source, synchrotron radiation only consists of the continuous background since the electrons are emitting the radiation as they are being accelerated by a magnetic field perpendicular to the electron orbit.

The fact that the electrons are relativistic --- they move, depending on the machine energy, at >99.99% of the speed of light --- gives rise to some important differences in the emission characteristics. These specific characteristics are depicted in Fig. 4.

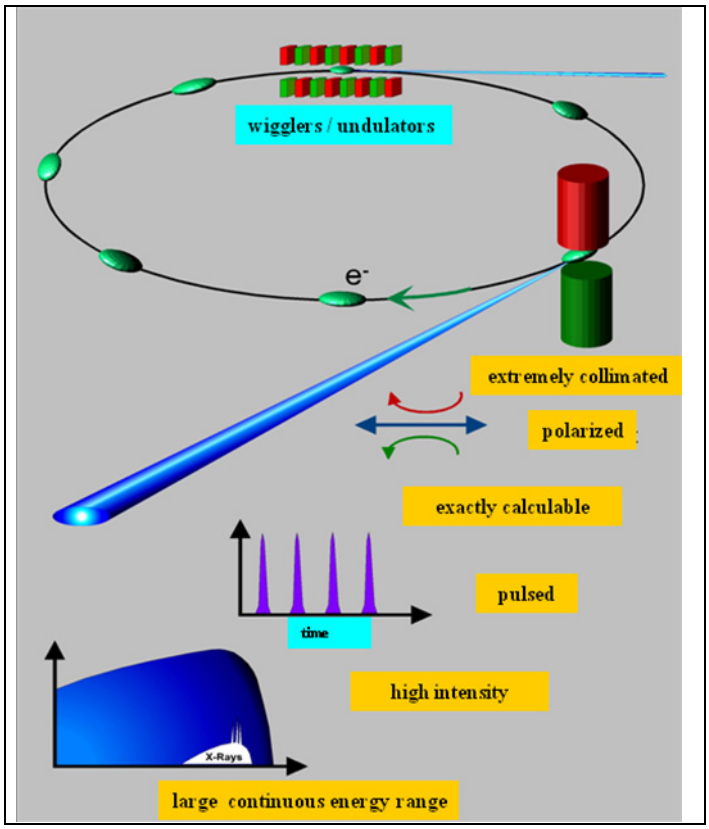


Fig. 4: *Properties of synchrotron radiation (source BESSY)*

2.1 Continuous spectral distribution

One of the most widely appreciated features of synchrotron radiation is the continuous spectrum from the RF and THz region up to hard X-rays. The normalized spectral distribution $S(\Omega/\Omega_c)$ of synchrotron radiation is shown in Fig.5. This function is universal to all sources and depends only on a few specific source parameters such as the electron beam energy and the magnetic field, which is sometimes also expressed by the bending radius of the deflecting magnet. This universal spectral distribution is actually a complex integral over some Bessel functions, the exact details are given in 1 and 2. It is characterized by a single parameter (Ω_c) whereby $h\Omega_c$ is the so called critical energy (E_c). The critical energy is defined as the energy, where exactly half of the power is radiated at frequencies below this energy and half at higher frequencies. The peak of the spectral distribution is actually located at about $\Omega_c/3$. It is also useful to note that the function $S(\Omega/\Omega_c)$ can actually be quite well approximated both for frequencies smaller and larger than Ω_c as indicated in Fig.5.

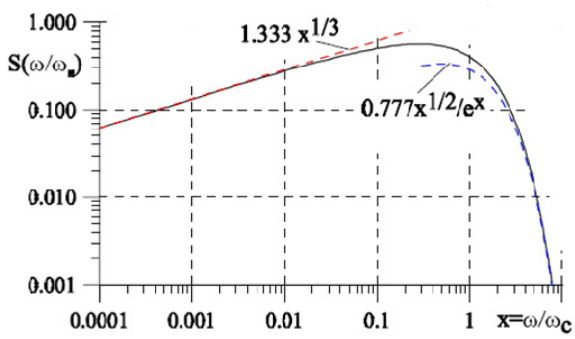


Fig. 5: *Spectral distribution of Synchrotron Radiation (rom ref. 1)*

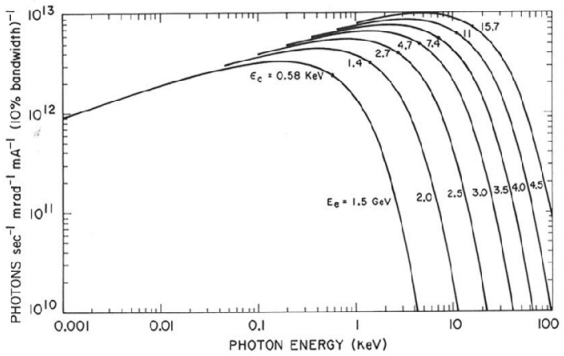


Fig. 6: *Spectral distribution as function of electron beam energy (from ref. 3)*

The **critical energy E_c** of a synchrotron source is given by

$$E_c = (6\pi c/h)\,(1/R)\,(E_e/E_o)^3$$

It is important to note that the critical energy depends to the third power on the electron beam energy (E_e/E_o) but only linearly on the magnetic field B (or inversely on the radius of curvature R of the particle in the magnetic field B). Since the bending radius R is proportional to Ee/B for application purposes one can transform this into

$$E_c\,(\text{keV}) = 0.665\,E_e^2\,(\text{GeV})\,B(\text{T})$$

Thus the characteristic energy can be related to readily available parameters. Typically bending magnet fields are in the order of 1T. Accordingly for a ring of 3 GeV electron energy, which is a standard parameter for modern storage ring sources, the critical photon energy is about 6 keV. The useful photon energy range extends up to about $5E_c$, thus such a source will be available for experiments requiring photon energies up to 30 keV.

Another extremely useful formula is the relationship between the photon energy (E_{phot}) and the wavelength λ as given by

$$E_{phot} = hc/\lambda$$

Or in practical units numerically

$$E_{phot}\,(\text{eV})\,\lambda\,(\text{nm}) = 1239.8$$

The power radiated by a single electron is given by

$$P_S = (e^2c/6\pi\varepsilon_o)(1/R^2)(E_e/E_o)^4$$

Apart from the fundamental constants (e,c,π,ε_o) the radiated power only depends on the electron energy ($\mathbf{E_e}$) and the radius of curvature of the trajectory (**R**) induced by the magnetic bending field. The scaling with rest energy of the particles $(E_e/E_o)^4$ is responsible for the fact that for practical purposes only electrons or positrons are used in synchrotron radiation facilities.

In practical applications the interest is in the number of photons emitted by the source and thus the beam current I enters linearly into the equations. Some actual curves showing the photon flux available at a storage ring source are shown in Fig. 6. These curves are scaled to a beam current of 100 mA, which is a typical value. The different curves nicely illustrate that below the critical energy the available photon flux at a given bending magnet radius does not vary with the electron energy in the ring, whereas the high energy cutoff for experiments depends critically on it. What also enters is the solid angle around the circumference of the ring, over which the radiation can be collected for the experiments.

2.2 Collimation and brilliance

Compared to many other sources the synchrotron radiation is **extremely collimated** in the plane of the accelerator. This is a relativistic effect and illustrated in Fig. 7. In the rest frame of the electron the radiation is emitted in the characteristic (donut shaped) dipole emission pattern. Since the electron is moving close to the speed of light howver this emission pattern looks totally different in the laboratory frame where the radiation is observed. The coordinate in the direction of motion of the electron is transformed via a Lorentz transformation, while the other two coordinates are not affected. Accordingly the emission is peaked very strongly in a narrow forward cone as indicated on the right hand side of Fig. 7. From any point along the trajectory, the light is collimated as shown in Fig. 7 and since the electron moves along a circular arc, the radiation is observable in tangential direction around the ring from all points of the electron trajectory, where the electrons are accelerated. If a single electron moves around a circular storage ring the light will be emitted in short flashes similar to the pulsed characteristics of a beacon or a light tower.

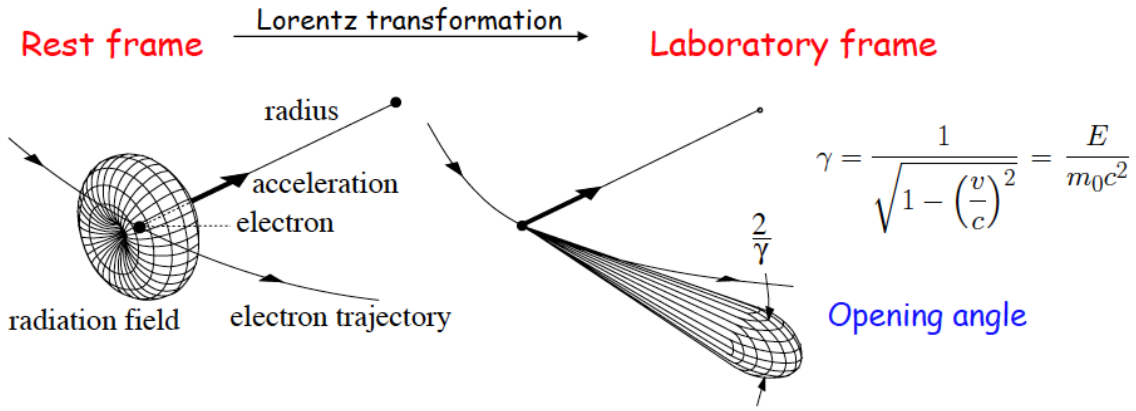


Fig. 7: *Emission pattern of the (dipole) radiation of a relativistic charged particle within its own frame of motion (left side) and as seen from an observer in the laboratory right side)*

The high collimation immediately relates to the **brilliance** of this light source. The brilliance is defined as the number of photons emitted within a certain wavelength interval per unit of source size and per unit of solid angle. Obviously here the intrinsic collimation is very favorable. Brilliance is the factor that distinguishes a laser from an ordinary light bulb. The source size of synchrotron radiation is given by the cross section of the electron beam, the emittance, and the spectral distribution and intensity as well as the emission characteristics are described above.

Apart from the continuous spectral distribution synchrotron radiation is much closer in characteristics to a laser than a light bulb. Furthermore, the brilliance is an intrinsic quality factor of any light source since the brilliance cannot be improved by an optical manipulation. Poor optics will only decrease the brilliance of the light source.

2.3 Intrinsic time structure

In an RF accelerator electrons can only be accelerated around a certain phase point of the RF field. Accordingly the electrons are bunched in packets of a certain length, depending on the phase space available for stable acceleration conditions. Typically such an electron bunch is about 50 ps long (1.5 cm) and the bunches are 2 ns apart (60 cm). Given the collimated tangential emission characteristics and the limited angle over which the electron trajectory is visible for the observer, this results in the time during which the radiation is observed.

2.4 Polarization

Synchrotron radiation is polarized. The polarization is linear in the plane of the accelerator and elliptical with different helicity, if viewed from above or below the plane. This is illustrated in Fig. 8. Viewed from the side in the direction of the peak of the emission cone, the charge is accelerated back ad forth in the plane and this acauses the electrical field amplitude to oszillate horizontally, parallel to the acceleration. Similarly, the elliptical polarization can be plausibly derived as indicated in Fig. 8, if the observer is located above or below the plane of acceleration. Analytically this can be derived from momentum conservation. In the emission process the electron looses a small amout of energy, and thus also (relativistic) momentum. For the combined system ---electron and photon--- energy and momentum together have to be conserved. Accordingly the photon carries a momentum corresponding to the momentum loss of the electron.
Incidentally, the polarization of the X-rays in space interstellar space is taken as strong evidence that these X-rays are indeed due to synchrotron radiation, produced when relativistic electrons are accelerated by the strong magnetic fields of stars, for example pulsars.

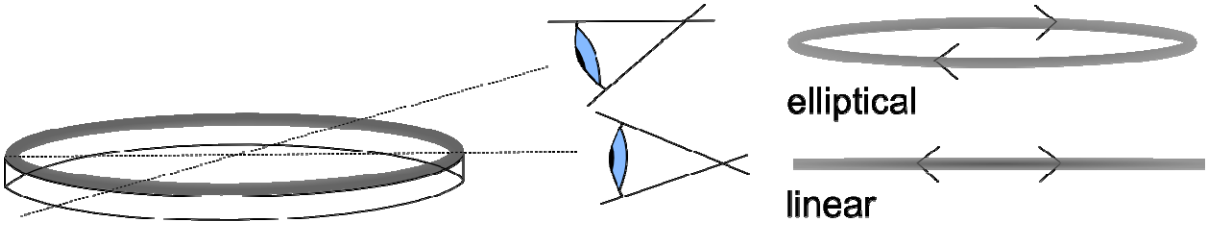


Fig. 8: *Sketch of a circular electron orbit to illustrate the polarization of synchrotron radiation (Source DESY)*

Obviously the angular distribution of both polarization components is different. While the parallel polarization component is strongest in the plane of the ring, the vertical polarization component vanishes in this plane and has a wavelength dependen maximum above and belo this plane as shown in Fig. 9. The exact analysis (1,3) also shows that the opening angle is wavelength dependent and Fig. 9 shows several distribution curves given for wavelengths λ in units of the critical wavelength λ_c.

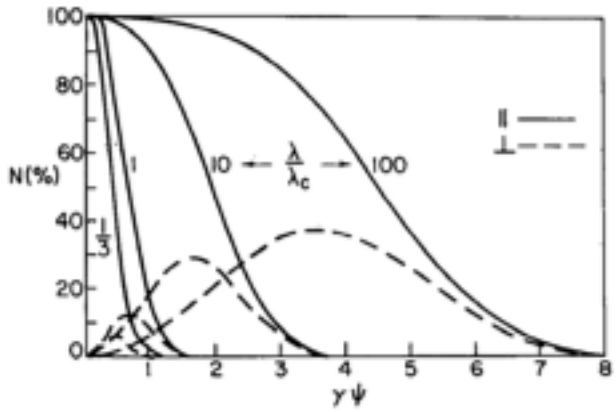


Fig. 9: *Vertical distribution of the two different polarization components for different wavelengths (in units of l/lc) (from ref. 1)*

2.5 Coherence

The coherence of a light source is another special feature describing the accuracy of the wavefront available for the investigations. The coherence enables important techniques such as holography or lenseless imaging. For a coherent light source the electrical field amplitude and phase is given at any point and time in space. As described graphically in Fig. 10 any light source can be made coherent by the appropriate measures. Fig. 10 a shows an incoherent light source emitting two different wavelengths (red and blue) into all directions from an extended source area. Placing an aperture in front of this source defines the source volume and also the wavefront in relation to this aperture. The curvature of the spherical wavefront originating from this aperture defines the lateral coherence length at a given distance from the aperture, where the wavefront is defined better than a factor of $\pi/2$. This is also called the spatial or lateral coherence. However the absolute electric field is still not well defined in time, as the two emitted wavelengths have differing oscillation periods.

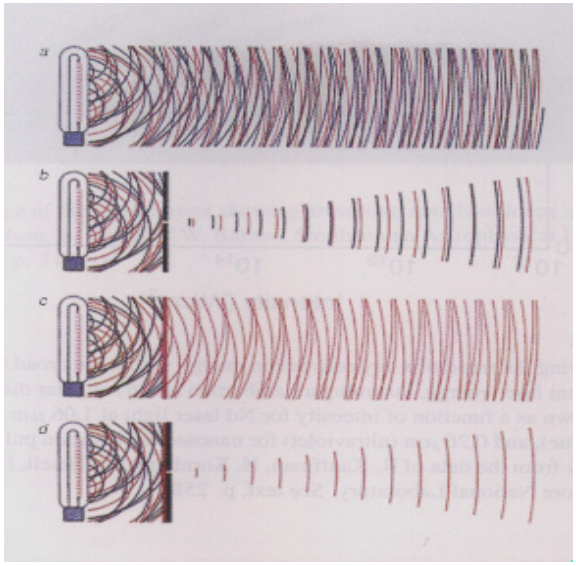


Fig. 10: *Improving the coherence of a light source (A. Schawlow, Sci. Am. 219, 120(1968))*

Fig. 10c illustrates the effect of placing a wavelenght selective filter in front of the light source. Now only the red light passes through. The combination of both measures, the aperture and the filter, now finally results in a coherent wavefront (10d).

The wavelength purity of a source is also synonamous with the temporal or longitudinal coherence. As just mentioned, any light source can be made coherent applying these measures, however the difference between a light bulb and a laser is how much intensity is left afterwards. The high collimation and small source size of the synchrotron radiation are very beneficial for making the radiation coherent. A couple of appropriately chosen apertures and a monochromator are all that is required. Since both the source size and collimation also define the brilliance of the source, the coherent flux available form a synchrotron scales directly with the brilliance.

2.6 Improving the performance: wigglers and undulators

Following the construction of dedicated synchrotron radiation facilities, ideas were generated how to improve the characteristics of a given facility. This is how wigglers and undulators came into play. Wigglers and undulators are periodic magnet structures constructed to compensate the overall deflection of the electron beam by alternating the direction of the magnetic field. These periodic magnet structures can be placed into a straight section of the accelerator or storage ring, as such they a referred to by the general term of insertion devices. Such an arrangement of magnets is shown schematically and in the real world implementation in Fig. 11. The difference between these two devices is only of a quantified nature. In a wiggler the magnetic field strength is such that the electron deflection angle in each pole is larger than the natural opening angle of the synchrotron radiation emission, whereas in an undulator the deflection angle of the electrons is smaller than the opening angle. Consequently in an undulator the entire electron trajectory is located within the cone of the emitted radiation. Obviously there is a smooth and continuous transition between these two regimes.

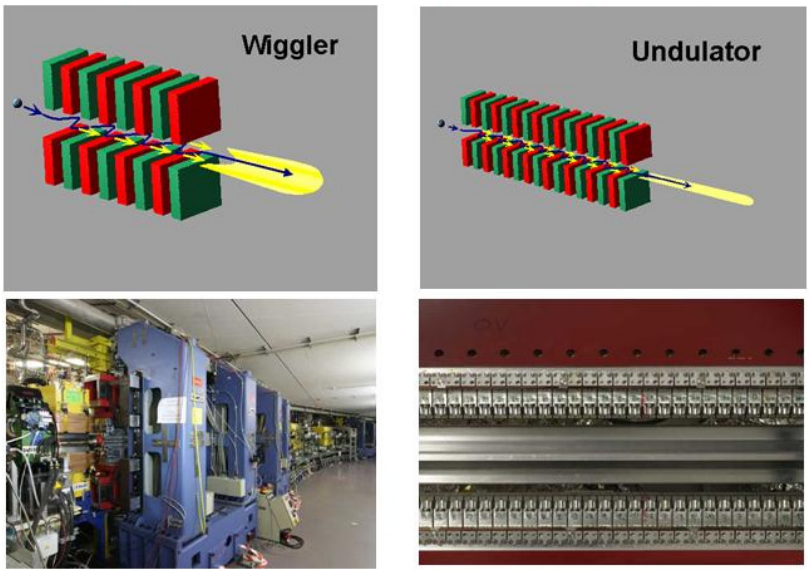


Fig. 11: *Top panels illustrate the periodic magnet arrangements in a wiggler or undulator, the bottom panel on the left shows an undulator as installed at the BESSY II storage ring and the bottom right panel is a view of the magnet assembly of this undulator (source BESSY)*

This seemingly small difference results in remarkably large differences in the characteristics of the emitted radiation. In a wiggler, or whenever the deflection angle of the electrons is larger than the natural collimation angle of the emitted radiation, the radiation intensity just is summed up incoherently along the electron trajectory visible to the experiment along the length of the device.

The main purpose of the wiggler however is to extend the usable photon energy range to higher energies. This is accomplished by wigglers having a larger magnetic field and thus a higher critical energy than the bending magnets of the same facility. As illustrated above in Fig. 6, the usable wavelength range at any given facility can thus be extended significantly to higher X-ray energies. Typically wigglers are constructed as a series of superconducting magnet poles, which have magnetic fields as high as 7T, whereas a typical bending magnet has a field of 1 to 1.2 T.

Analytically these devices are distinguished via the dimensionless deflection parameter K.

$K = \alpha\gamma$

Where α is the maximum angle of deflection of the electron beam in the device and. Given the maximum strength of the (sinusoidal) magnetic field B_o and the period of the undulator (wiggler) λ_u, K can be calculated as

$K = 0.934\ \lambda_u\ (cm) B_o(T)$

For K > 1 obviously the deflection angle is larger than the opening angle of the radiation and this is a wiggler, whereas for K< 1 the entire trajectory is contained within the natural radiation cone of the synchrotron radiation. The angular emission characteristics of the wiggler is accordingly enhanced by the deflection angle to $K\gamma$ horizontally, while it remains confined vertically to $1/\gamma$ as shown in Fig. 12 left panel.

In the undulator the electron and the photons travel essentially at the speed of light, but the electron has a slightly longer path, since it travels along a curved (sinusoidal) trajectory. This leads to a coherent superposition of the emission of the same individual electron, when the pathlength of the electron and photon differ by the wavelength of the emitted light.for each undulator period. Such the emission spectrum of the undulator is characterized by strong harmonics peaking at wavelengths λ_n given by

$\lambda_n = \lambda_u/2n\gamma^2\ (1 + K^2/2 + \gamma^2\theta^2)$ for n = 1,2,3,4……

where n is the number of the harmonic and θ is the angle of the observer relative to the axis of the undulator. Using the standard equivalents this can also be written in terms of the photon energy of the harmonics E_n as

$E_n\ (keV) = 0.95\ n\ E^2\ (GeV)/\lambda_u(cm)(1 + K^2/2 + \gamma^2\theta^2)$

The opening angle of the emitted radiation also is modified by taking the number of periods N of the device into account as shown in the right hand panel of Fig. 12. Typically undulators are constructed from blocks of permanent magnets, a design originally suggested by K. Halbach (4) from Berkeley. This allows for a very precise assembly of the magnetic structures.
Mechanically opening and closing the gap between these magnet assemblies, changes the magnetic field and thus the undulator parameter K and spectral output. The necessary mechanical precision for such a motion is enabled by the massive mechanical structure holding the magnet assemblies in place as seen in Fig. 11.

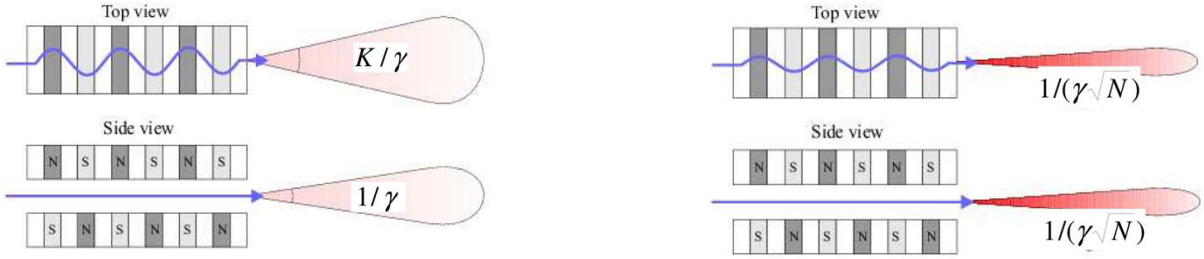


Fig. 12: *Trajectory and emission cone in a wiggler (left side) and in an undulator (right panel) (source DESY)*

The modification of the spectral output at the same synchrotron source (SPRING-8, Japan) by insertion devices is illustrated in Fig. 13. The wiggler extends the bending magnet spectrum to higher X-ray energies and the flux (and brilliance) adds up proportional to the total number of periods N. The undulator has a peaked spectrum, where the peak brilliance scales as N^2. In order to produce different wavelength outputs from an undulator this device needs to be tuned by changing the magnetic field strength. This is conveniently done by changing the vertical gap between the poles of the magnet structure. Stronger magnetic fields cause a stronger deflection of the electrons and thereforte a longer pathlength for the eelctrons and accordingly the energy of the corresponding harmonics is lower. The total tuning range for the 1st, 3rd, and 5th harmonics is also indicated in Fig. 13 Pulling the magnets apart drives the energy of the harmonic up, but this lowers the critical energy of the emitted radiation. Consequently the brilliance (and intensity) of the harmonics exhibit a sharp drop towards higher photon energies.

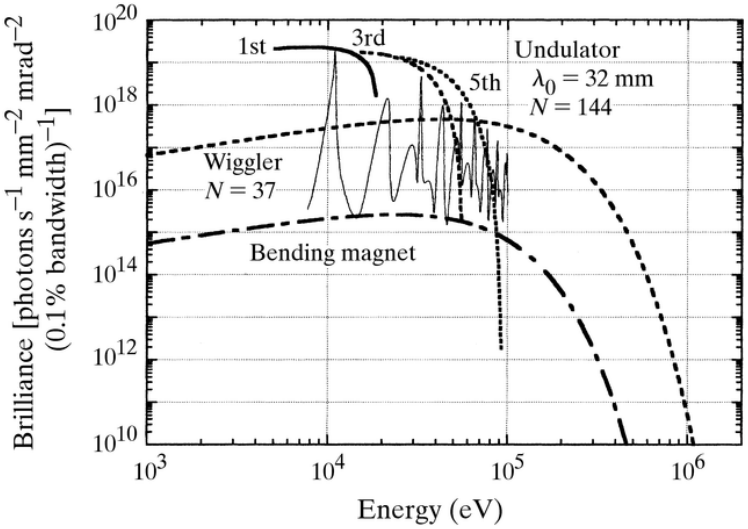


Fig. 13: *Spectral brilliance of various sources (bending magnet, wiggler, undulator) at the SPRING-8 facility (Japan) (from SPRING-8, IUCR 1998)*

Undulators also offer a unique control of the polarization of the radiation. This is achieved in the so called APPLE devices, where a spiraling component can be impressed upon the electron trajectory. In the most advanced APPLE-II devices the electron trajectory can be completely controlled such that circular polarization of selectable helicity as well as linear polarization with a selectable orientation of the polarization plane can be produced (5).

3. State of the Art and Future Sources

3.1 Brief History of synchrotron radiation sources

Following the initial discovery of synchrotron radiation, the light was initially used on various electron acclerators where it is produced as an unwanted by-product, which actually limits theultimately achievable energy of the synchrotron. This so-called parasitic mode of operation was carried out at a few laboratories in the world. Some of the earliest beamlines and laboratories were located at NIST (Gaithersburgh, USA), Daresbury (UK), Frascati (IT), and DESY (Hamburg). Compared to today these operations were carried out under much less favorable conditions. Nevertheless, since the field was very new, quite a few pioneering studies were performed at these sources, which paved the way for the growth of the community and better conditions. These were then offered at the so called second generation sources, which were actually dedicated exclusively to the production of light. The first facility of this kind in the world was the tiny storage ring Tantalus in Wisconsin (USA). Now the emphasis could be placed on stable and reproducible conditions and the field of synchrotron radiation research gained a large momentum. Dedicated facilities were constructed all around the world. The next and present state of the evolution are the third generation facilities, where the emphasis is on exploiting the special performance of undulator based sources. These fall into two categories, the high energy facilities -----ESRF (Grenoble, France), APS in Chicago (USA) and SPRING8 in Japan---- where the performance is optimized for the hard X-ray range and low energy facilities, such as ALS Berkeley (USA), Maxlab (Lund, Sweden), SLS (Switzerland), BESSY II, Elettra in Trieste, Italy where the peak performance is in the VUV-Soft X-ray range. For a complete list of all facilities see (www.lightsources.org). The next and fourth step of the evolution is the development of free electron lasers, which again offer a dramatic increase in source brightness over the previous undulator based sources. Even though the source characteristics are quite different from the other more conventional sources, (see the description in the following chapter) these are regarded as the fourth generation of synchrotron radiation sources.

3.2 Storage Rings

Storage ring sources form the basis of the success of accelerator based light sources over the last decades. They have reached a level of technical sophistication that they are operating extremely reliably and reproducibly close to the theoretically calculated performance limits. This is manifested by the use of storage rings as (legal) calibration standards for electromagnetic radiation over the wavelength range from THz to X-rays. Third generation synchrotron radiation sources based upon high brilliance storage rings equipped with a multitude of specialized insertion devices for the ultimate photon source performance are serving an ever increasing community of scientists and science disciplines. Worldwide more than 50 facilities are operational for research, metrology, and fabrication with ten's of thousands of researchers performing their individual experiments at a highly competitive level. 12 of these facilities belong to the most modern 3rd generation. Several more facilities are under construction worldwide, but the demand for beamtime still outpaces the development of new capabilities.

The **wavelength range** of storage rings covers the full spectrum from **THz-radiation to hard X-rays** with selectable and precisely controllable **polarization** of the light produced in undulators and wigglers installed in straight sections of the storage ring. The pulse length is typically several 10's of ps, even though special electron optics schemes or fs slicing with an external laser have been incorporated at several facilities allowing for studies with

substantially reduced pulse lengths down to sub 100 fs, even though at a considerably reduced intensity. State of the art facilities have 50 or more experiment stations, where research can be performed with individually and independently selectable photon characteristics as best matched to the requirements of the experiments.

As with all accelerator based sources, the performance is based upon a combination of properties of the electron beam as well as the magnetic devices used to generate the radiation. The **emittance** ε is the intrinsic parameter describing the source size and divergence of the electron beam and thus the intrinsic quality factor of the electron beam. The emittance of storage rings scales as E^2/R where E is the electron energy and R the bending radius. Superimposed on the electron beam emittance are the characteristics of the (insertion) device used to produce the photon beam. The combination of these parameters determines the **brilliance** of the photon beam. Modern storage ring facilities operate in a parameter range where the electron beam emittance contributes significantly to the brilliance of the photon beam.

The **brilliance** of the photon beam, as the **emittance** of the electron beam, is an intrinsic quality factor. The brilliance is not altered by focusing (Liouville's theorem) and essentially limits the number of photons that can be focused into a certain spot size. Imperfect optical components used to manipulate the photon beam obviously deteriorate the brilliance. Thus any kind of microscopy/microanalytics experiment immediately benefits proportional to the improvement in the brilliance, whereas many conventional experimental techniques scale in performance proportional to the total **photon flux** arriving at the experiment. The coherence of the photon beam also scales with the brilliance.

Typical parameters of storage ring sources concerning the (average) brilliance are shown in Fig. 14. Here European sources are taken as examples. The US sources, including the ones under construction such as NSLS-II, exhibit a comparable performance.

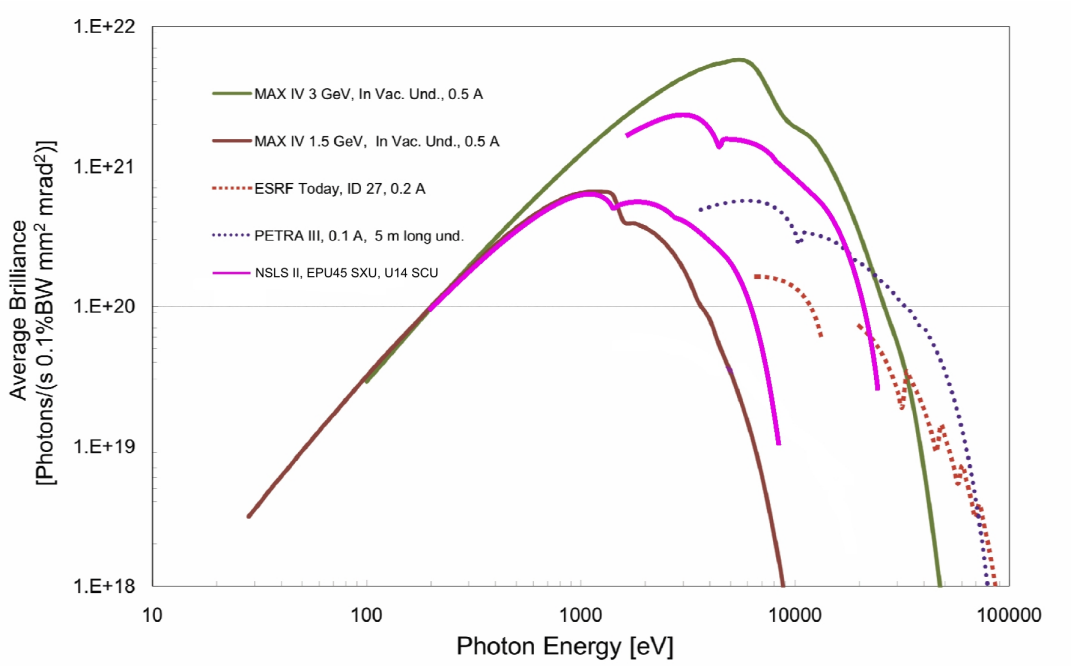


Fig. 14: *Average brilliance of the photon beams generated by various insertion devices at existing (ESRF, PETRA III) storage ring sources including calculations of the NSLS-II and the planned MAX IV facility. (6)*

Obviously other photon beam properties are of relevance to the experiments as well. These include for example the **polarization**, which can be freely selected between circular and linear polarization and even the plane of the linear polarization can be arbitrarily rotated from horizontal to vertical using state of the art undulators. Furthermore at storage rings the complete **photon energy range** is accessible from **THz to hard X-rays** without any change in the electron beam parameters. The details of the performance and the individual tuning range depend on the specific insertion device and the corresponding beamline. **Tuning** the photon energy (and resolution) typically is performed on the timescale of seconds. The **energy resolution** can be as high as $E/\Delta E = 10^6$ depending on the monochromator or optics used. In special cases even higher resolution values are achievable.

Time resolution is typically limited to 20 to 50 ps, depending on the length of the electron bunch circulating in the storage ring. Special electron optics allow to reduce the bunchlength to produce photon pulses in the sub ps range, even though at a reduction of overall bunch charge (beam current). Slicing the electron bunch with an external laser was pioneered at the ALS (7) and has been implemented also at BESSY-II (8) and the SLS, giving sub 100 fs pulses in the soft to hard X-ray range, again at a reduced overall intensity. Contrary to laser based (HHG) sources these pulses can be completely controlled with respect to the polarization parameters and thus cannot be produced by any laboratory (table top) laser scheme in a similar quality.

The **coherence** length of the photon beam is on the order of 10 to 30 μm, both longitudinally and laterally. The coherence directly scales with the brilliance of the beam and decreases with increasing photon energy. As one important example of experiments, coherence enables holographic imaging techniques to be exploited under careful consideration of the photon beam properties.

The question arises about the performance of the 'ultimate' storage ring facility. Theoretically this can be approached by setting the electron beam emittance to zero in the calculated performance curves. Even though this seems an academic exercise, it nevertheless defines the intrinsic limitations for all accelerator based light sources. Fig. 15 shows the theoretically possible improvement of the photon beam brilliance for some examples while all other parameters are kept fixed. From these graphs one can deduce that for example around 1 keV photon energy for a 3 GeV electron beam an improvement by one order of magnitude can be expected, while for a 7 GeV electron beam this improvement is noticeably larger. Also at higher photon energies, around 10 keV, the possible improvement using an 'ultimate' storage ring as compared to a state of the art present day facility is more than 2 orders of magnitude. This is in line with the above reference scaling of the electron beam emittance proportional to the E^2/R. A'real' storage ring at a larger energy will intrinsically have a larger emittance and thus produce a less brilliant photon beam.

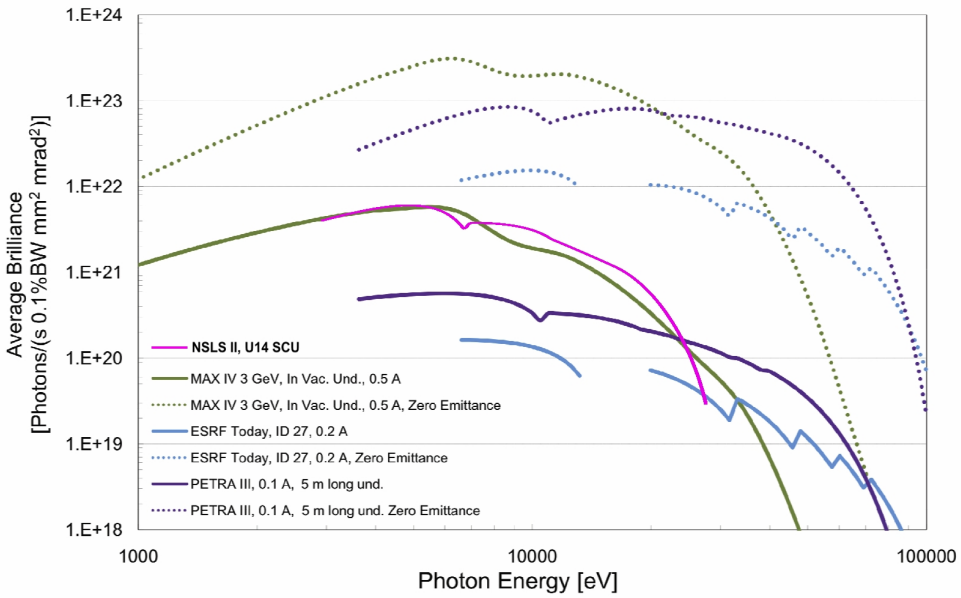


Fig. 15: *Average brilliance of several insertion devices at state of the art storage ring facilities and the influence of the electron beam emittance. The finite emittance reduces the average brilliance from 1 to 3 orders of magnitude. From this comparison the performance of these same ID's in an ultimate storage ring can be estimated. (6)*

3.3 Linac Based Photon Sources: Advancing beyond third generation sources

The synchrotron user community continues to grow steadily, attracting for example new science fields such as archeology and cultural heritage, and even more important, the demands regarding the quality of the photon source parameters are also increasing noticeably. The experiments get more demanding and sophisticated. The analysis of nanoscale systems, with respect to structure and electronic properties, has generated a large increase in interest from the user community not only in materials sciences, but also in biology as well as life- and environmental sciences. Individual nanosized structures and materials tailored to deliver novel dedicated performance require an ever increasing spatial resolving power of our analytical probes. The implications for the light source design are straightforward. Lower emittance electron beams result in higher brilliance photon beams, as well as a higher coherence fraction, and thus directly improve focusing on the nm scale for high resolution microscopy and spectroscopy applications.

Additionally, as the experience at the FLASH facility and at LCLS demonstrates, there is a substantial interest to study dynamic processes requiring high-brilliance, *fs* X-ray pulses which are generally outside the range accessible by storage rings. Proposals for experiments range from charge carrier dynamics in materials, to chemical reaction kinetics. Another interesting question relates to the development of an understanding of the processes determining the ultimate timescale of magnetic data recording. Initial pioneering studies of this kind have been performed at femto-slicing facilities in storage rings (7,8), but only with great effort on account of the low flux. Furthermore, not only shortest pulse lengths are important, but also the ability to control the exact shape of the pulses and to synchronize these to external (laser) sources.

The realization of many of these sophisticated and demanding investigations are clearly beyond the capabilities of existing storage ring based sources. Accordingly one has to use a

different approach. The next-generation light sources must circumvent the equilibrium beam properties of a storage ring source, i.e., the beam must be discarded before it has a chance to sample and occupy all of the phase space offered by the magnetic lattice of the storage ring. A single-pass linear accelerator represents the most extreme example of such a device. Contrary to the case of storage rings, the electron beam emittance decreases in a linear accelerator as the energy of the electrons increases. Provided one takes care not to disrupt the beam, it maintains the properties generated in the linac, with the emittance being *damped* adiabatically by a factor inversely proportional to the beam energy. Ultra-low emittances are thus attainable by acceleration provided the electron source produces a beam of suitable quality.

Other parameters, especially the bunch length, can be optimized by bunch compression techniques along the linac, providing a high degree of flexibility to address the user requirements. Sub-ps pulses can be generated, up to 1000 × shorter than in a storage ring. These operational modes of a linear accelerator driven light source have already resulted in pulse lengths on the order of about 10 fs at the FLASH facility in Hamburg.

Two different types of linear accelerator based light sources have to be discussed here:
Free Electron Lasers (FEL) and **Energy Recovery Linac (ERL)** sources. Even though both use ultra bright linac generated electron beams to produce the photons, their source characteristics are sufficiently different, such that they serve different purposes and largely different science areas. The typical layout of an FEL is shown in Fig. 16a and the generic design of an ERL is shown in Fig. 16b.

In the FEL the electron bunches are accelerated in the linac and then distributed alternately into several long undulators to produce FEL radiation at the individual wavelength as given by the beam energy and the undulator (ID) parameters. After the FEL process the emittance of the electron bunch has deteriorated such that it cannot be used for the FEL process again (or only at substantially longer wavelengths) and the beam is dumped. Accordingly the FEL serves only a few experiments at any time (typically between 1 and 10) with photon pulses of extremely high peak brilliance. Whenever experiments are considered requiring a high average brilliance one has to take into consideration the number of experiments that are served by the same linac in parallel.

The ERL, on the other hand is a true multi user machine. The ultra bright electron bunches are accelerated and then passed through an arc of insertion devices, very similar in layout as a storage ring. At the end of the turn, the bunch is passed into the linac again, to decelerate the electrons and to regain the beam energy. This substantially helps to reduce the operational costs of the facility. The important characteristics of the ERL however is that the electron bunches maintain the properties of the linac source and never reach the equilibrium properties of a bunch circulating in a storage ring. It typically requires thousands of turns in a storage ring for a freshly injected bunch to reach equilibrium state. Thus the ERL can be viewed as the successor of the storage ring, whereby all (50 or so) users share the same bunches, but the ERL excels concerning the average brilliance because of the lower emittance of the electron beam.

For the production of photons in the IR range, both FEL facilities and ERL facilities exist and have generated a well documented scientific impact. The challenge arises from the production of shorter wavelength photons. The required performance of these accelerator driven sources is based upon the emittance of the electron beam that has to scale relative to the wavelength of the light to be generated.

Accordingly, an FEL or ERL operating in the X-ray range needs to have an electron beam emittance which is at least factor of 10^3 smaller than an ERL or FEL operating for the production of IR photons.

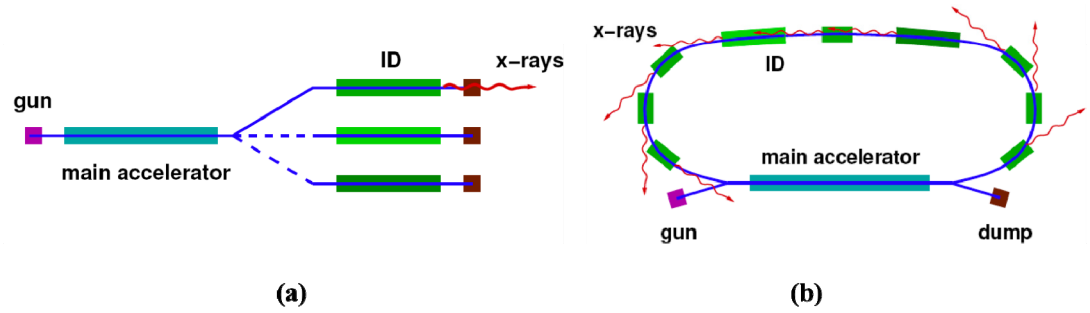


Fig. 16: *Schematic layout of an FEL facility (a) and an ERL facility (b).*

3.4 Free Electron Lasers

Free electron lasers offer an outstanding performance as far as the peak brilliance of the photon pulse is concerned. They will enable the study of non-linear phenomena and processes or the investigation of very dilute samples, such as individual sizes selected nanoparticles or biological molecules or highly excited short lived species in plasma processes. The combination with a second, external laser source, either used to excite or create a certain species offers unique prospects for science.

The basis for the special performance of the FEL is due to an internal modulation of the electron density in the bunches. Once this modulation has been achieved at a certain wavelength λ, all the electrons in the bunch will emit coherently at this wavelength. Thus the electron beam has been transformed into a laser medium. This electron density modulation can be achieved by seeding with an external laser or by spontaneous self modulation. This **SASE** (Self Amplified Spontaneous Emission) process, is illustrated in Fig 17. As the electron bunch travels through a suitably long undulator it spontaneously emits radiation, largely enhanced at the wavelength given by the harmonics of the undulator radiation. The interaction of this photon field with the electron bunch results in an energy transfer and subsequently a spatial modulation of the electron density in the bunch at the wavelength of the emitted radiation. Once this spatial modulation is achieved, all electrons emit coherently. This results in a dramatic gain of the radiation emitted at the principal wavelength of the undulator and at higher harmonics. Eventually the gain saturates, since the spatial modulation of the electron bunch is destroyed by the emission of the photons.

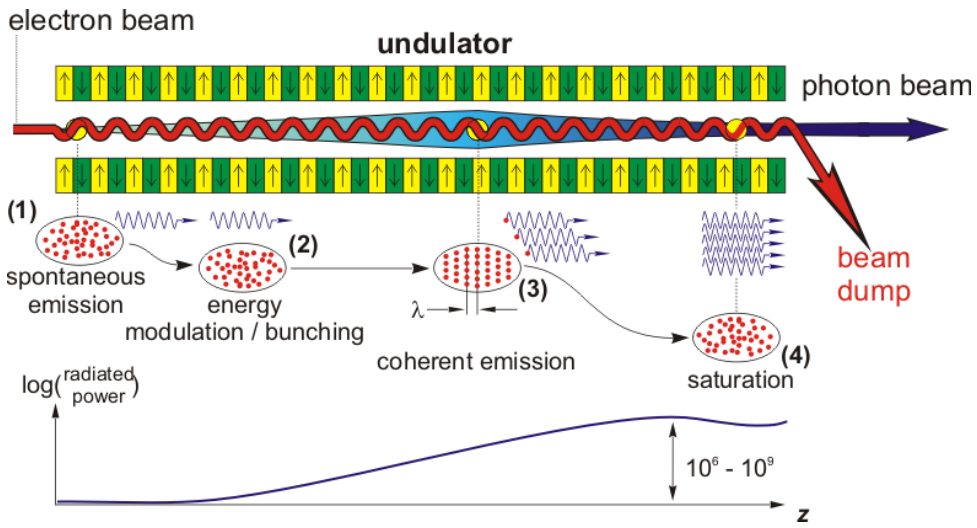


Fig. 17: *Schematic of the SASE process. The electron bunch is modulated in space by the interaction with the photon field it generates. The bottom part shows the gain curve for the SASE FEL emisson (source DESY)*

This coherent emission process scales proportional to n^2 , where n is the number of electrons in the bunch contributing to the process. At 1 nC bunch charge, which is the canonical number for any FEL project, this corresponds to approximately a factor of 10^9 compared to the superposition of the incoherent emission of the same electrons. This dramatic improvement results in a peak brilliance, which is unrivalled by any other source in the X-ray range as shown in Fig 18. Accordingly the FEL is uniquely suited for experiments requiring a high **peak brilliance**. The **coherence** of the photon beam scales proportional to the peak brilliance. The degeneracy of the coherent photon pulse will be up to 10^7, meaning that there are 10^7 photons per coherence volume. Storage rings have a degeneracy parameter of 1 and at an ERL this might reach up to 10^3. This will enable new physics to be investigated. Moreover, any investigation involving non-linear processes clearly benefits directly from this peak brilliance, scaling proportional to the n-th power of the peak brilliance for an n-photon process.

The polarization of the FEL radiation is determined by the insertion device. Thus it is absolutely no problem to generate **circularly polarized (soft) X-rays**, a feat which is close to impossible using conventional laboratory laser sources.

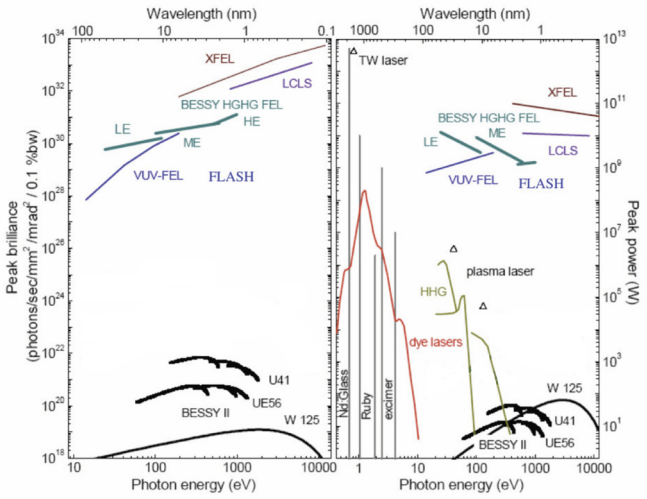


Fig. 18: Peak Brilliance and Peak power of several existing or planned FEL facilities compared with storage ring undulators and laboratory laser sources. (source BESSY)

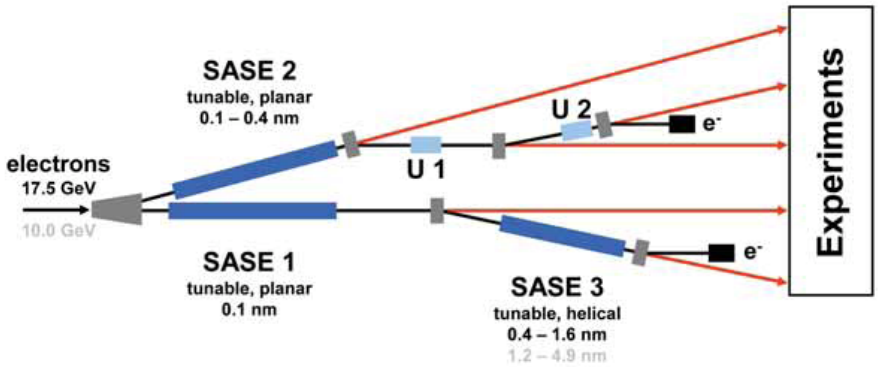


Fig. 19: Layout of the photon sources at the initial stage of the European X-FEL facility (9)

The layout of a typical FEL facility is shown in Fig.19 , which is taken from the TDR of the European X-FEL facility (9). The beam coming from the superconducting Linac is sent into either one of the two FEL branches labeled SASE 1 and SASE 2 respectively. After the bunch in SASE 2 has reached saturation, the emittance of the electron beam has grown substantially, such that the same bunch can only be used for normal radiation in the spent beam undulators U1 and U2. Similarly, after SASE 1, a SASE FEL process can only be reached at substantially longer wavelength (SASE 3), which is proportionally less stringent as far as the emittance is concerned.
Due to the large gain experienced in the coherent FEL process and despite of the low pulse repetition rate, the **average brilliance** of an FEL facility, is quite high, in the range of ERL sources, generally surpassing average brilliance values of todays storage rings. Most tabulated average brilliance values of FEL facilities published so far, however do not take into account that the accelerator beam will be distributed over up to 10 stations in an advanced FEL facility. Furthermore, the average brilliance has to be viewed with caution also to the extent, that for experiments requiring high average brilliance non-linear effects could be detrimental and special care has to be taken, that the high peak brilliance does not destroy the sample. This will be discussed in more detail in the section comparing the various facilities below.

The second outstanding feature of the FEL is that the **pulse lengths** are typically 100 fs or less. This enables time resolved investigations of a quality far beyond anything possible at storage ring sources. Time resolved diffraction or imaging enables the investigation of structural changes or phase transition in real time. Time resolved spectroscopy allows to monitor chemical reactions in real time or to investigate charge carrier generation, propagation, and collection processes in solar cells for example and to identify traps and recombination centers which deteriorate the device performance.

Pump probe investigations very often will require the **synchronization** of an external laser to the FEL pulse. Similar to synchronizing two independent table top lasers operating in the visible or IR range, jitter and drift of the synchronization will deteriorate the temporal resolution achieved in these studies. There is a substantial effort going on at various FEL facilities to develop schemes to reduce the jitter between these sources and thus to improve the synchronization. Moreover advanced seeding concepts will substantially reduce this problem as described below.

Pulse repetition rates of FEL pulses depend largely on the LINAC design used. Conventional (warm) Linacs are operating at a 10 to 120 Hz cycle. In each cycle up to several 10^3 electron bunches can be accelerated, spaced in time according to the frequency of the RF field. Thus the **pulse pattern** of an FEL has two characteristic frequencies. The macropulse frequency, is given by the main linac frequency between 10 and 120 Hz, while the individual pulses within each macropulse are spaced by 2 ns for a 500 MHz RF frequency. Superconducting linacs are being proposed for advanced FEL designs operating in a so called DC mode, where the RF field of the linac is permanently turned on. This will enable a more evenly spaced pulse pattern, for example tens of micropulses spaced at a repetition rate of (several) KHz. The limitations with respect to the overall number of pulses of an FEL currently are between 10^4 and 10^5 pulses/s. This limitation arises from the gun design of the accelerator complex.

The **energy resolution** $E/\Delta E$ of the pulses typically does not exceed 10^3. This is the inherent bandwidth of the emission from a long undulator. At best the pulse is transform limited and thus the pulse length and energy spread are coupled via the uncertainty limit. Monochromatization is possible at the expense of time resolution. Any grating will lengthen the pulses proportional to the number of grating lines used. Special compensating optics designs for monochromatization using two gratings have been suggested, but this will reduce the intensity by typically one order of magnitude for each grating.

Tuning the photon energy of the FEL is achieved by changing the gap of the undulators or the electron beam energy or both. Undulator gap tuning typically allows changes by a factor of 4 in photon energy or correspondingly in wavelength. This is comparable to the tuning range of a storage ring undulator in the first harmonic. Since the FEL is a high performance machine requiring utmost precision in terms of the emittance and compression of the electron bunch, tuning the FEL to a substantially changed photon energy or wavelength, is a lengthy and elaborate process. At presently operating facilities (FLASH) typically this takes hours to accomplish. This is very much the situation experienced with tunable fs lasers 1-1/2 decades ago. Developments of laser technology have resulted that by now this can be achieved at a turn of a button. (If no change of optics is required) With more experience and diagnostics it should be possible to tune the FEL wavelength at significantly shorter timescales than presently. Nevertheless, changing the photon energy at an FEL will always be much more delicate than at a storage ring source. Experiments requiring a rapid scan of photon energies over a larger range, such as EXAFS investigations, are not suited for FEL's.

2^{nd} generation FEL facilities; Seeding concepts

As the SASE process starts with the spontaneously emitted radiation in the first part of the undulator, the pulse characteristics are stochastic and intrinsically not reproducible. The wavelength and temporal structure will differ from pulse to pulse within certain boundaries, as different modes are contributing to the overall pulse envelope. This very much resembles the case of a short pulse laser without mode locking. Furthermore the synchronization to external lasers for pump-probe studies is difficult.

Accordingly ideas were generated to circumvent these shortcomings of the SASE process. It was demonstrated quite early at BNL that an external laser can be used to initiate the spatial modulation of the electron bunch, which prepares the 'medium', the electron bunch. for the coherent laser process (10). Thus the pulse properties of the external laser are imprinted onto the laser medium and the electron bunch then radiates with these pulse properties at the same frequency (wavelength) or at a higher harmonic of it. Essentially the FEL is used in this mode as a powerful amplifier of the externally produced photons.

This so called seeding process results in a much higher stability and reproducibility of the pulse shape of the FEL radiation and also has the benefit of the exact synchronization to the external seed laser pulse. A split off portion of this seed laser pulse can be used for the pump probe experiments, thus eliminating the dominant contributions to jitter between the two sources.

The quality improvements by seeding as compared to SASE radiation are shown in Fig.20

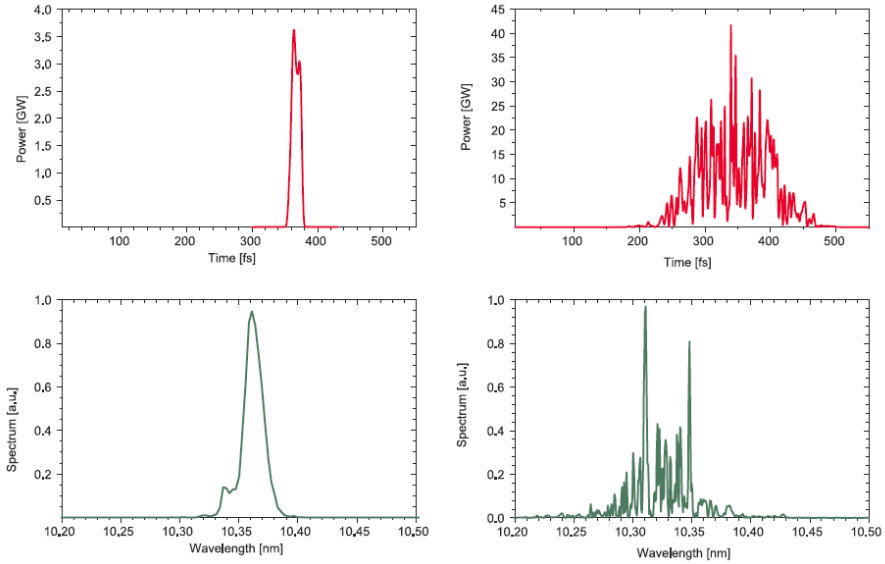


Fig. 20: *Model calculations of FEL pulses at 10.3 nm wavelength generated in a seeded (HGHG) FEL (left panels) in comparison with SASE-FEL pulses (right panels) (11)*

These are numerical simulations of a seeded FEL(left) and the corresponding SASE FEL on the right (11). Clearly the quality enhancement is visible, even though the peak power of the seeded FEL pulse is reduced by about one order of magnitude. The simple reason for this is that in the seeded FEL only the part of the electron bunch which is modulated by the seed laser contributes significantly to the FEL emission. The seed laser pulse can be generated for example by HHG, as described later. Alternately a cascading scheme can be used, where the seeding process is cascaded through several stages, always producing radiation at a higher harmonic (3^{rd} or 5^{th}) of the seed laser. This is referred to as HGHG (High Gain Harmonic Generation). The present consensus is, that these seeding schemes will enable to reach photon energies as high as 1 keV or wavelengths as low as 1.2 nm. For harder X-rays it is difficult to get a seed laser with sufficient power. Also the HGHG cascading is limited in the number of conversion stages, since with each stage the accuracy in reproducing the pulse shape of the seed pulse is not 100% perfect. Because of the obvious advantages of the seeding, most of the future low energy X-ray FEL facilities, planned or under construction, will implement seeding in one form or another.

The importance of seeding and online diagnostic tools for the verification of the pulse shape of the FEL is highlighted by the first actual measurements at the FLASH facility with an autocorrelation experiment (12). Fig. 21 shows the visibility of the two beam interference pattern using an autocorrelator to split, delay, and finally recombine the FLASH pulse. The visibility of interference at a delay of 50 fs directly proves that the FLASH FEL pulse has to have significant intensity over this time span. The pulse shape deduced from these measurements actually corresponds to a triple pulse structure as shown on the left of the figure.

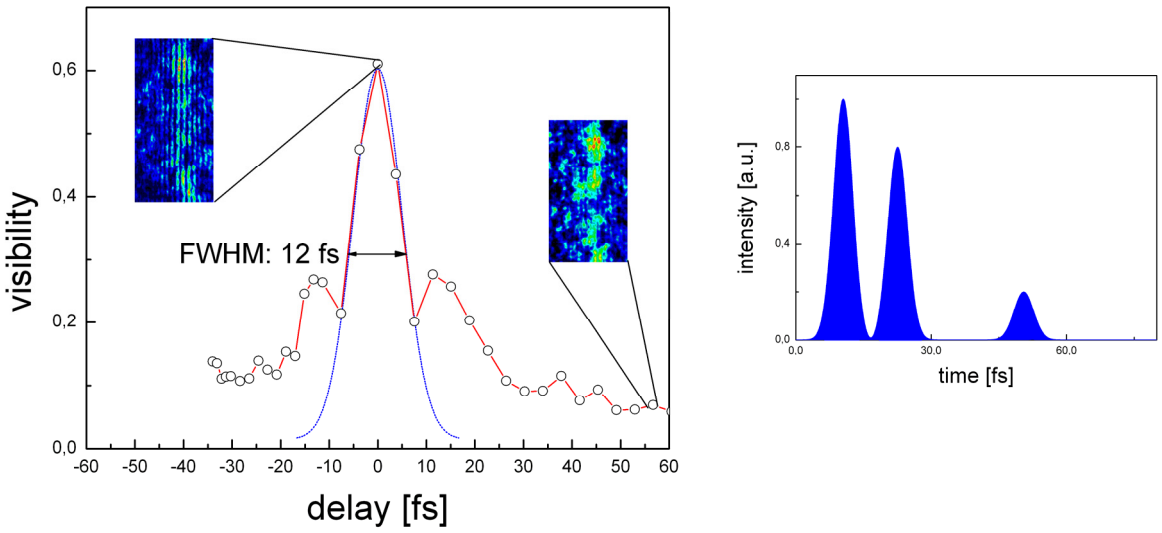


Fig. 21: *Linear autocorrelation of the FLASH-FEL pulse at 23.9 nm utilizing two beam interference. Plot of the experimental visibility of the spatial fringes and a Gaussian fit to the main peak (left) and pulse shape deduced from these measurements (right). This experiment provides proof of principle that pulse lengths in the 10 fs range can be achieved by FEL's (12)*

There is a simple way to discuss the implications for the science enabled by the control of the pulse shape or hampered by the lack of this control. If the experiment only uses one photon per event from the FEL it not ery much influenced by the exact temporal pulse shape. This is the case in one-photon events in dilute samples. For example, the determination of the structure of an individual nanoparticle by scattering does not depend on the pulse shape of the FEL pulse, unless the nanoparticle is destroyed within the timeframe of the FEL pulse. Spectroscopy on highly diluted species is another example which is possible, irrespective of the knowledge of the exact pulseshape.

On the other hand, if nonlinear (multiphoton) processes are investigated, then the exact pulseshape is of utmost importance. Using lasers in the visible or near UV, the pathway of for example a two photon photo-chemical reactions can be deliberately changed by changing the temporal shape of the photon pulse. Also for any pump-probe study, the temporal shape and synchronization of the photon pulse is of extreme importance.

In summary, FEL's are fantastic new photon sources in the (soft) X-ray range, enabling experiments previously impossible. These novel capabilities specifically extend to studies of nonlinear processes and phenomena and to time resolved studies on the fs timescale. Compared to storage rings a FEL facility has a substantially reduced number of experimental stations, which have to share the beam from the same accelerator. Therefore, FEL's probably never will replace existing storage ring facilities. Rather they should be viewed as complementary facilities that substantially broaden the horizon for science using X-ray photons.

3.5 ERL's: the 'storage' rings of the future?

While storage rings can accommodate a large number of different experiments over a wide photon energy range, they offer a high average flux, and are quite efficient to operate, they

also have certain limitations. Since the electron beam is at in equilibrium state, the emittance scales as E^2/R, requiring a very large R to reach small emittances. Furthermore, the energy spread of the electrons and the bunch length also deteriorate by the equilibration. The energy spread is typically > 10^{-3} and the bunch length between 20 and 50 ps. Sub ps bunches can only be produced at a substantial reduction in beam current, directly reducing the average brilliance.

Energy recovery linacs are accelerator based photon sources where the advantages of storage rings are combined with those of a linear accelerator driven photon source. Energy recovery linacs maintain the excellent beam properties of the linear accelerator, where the emittance scales as 1/E. The pulse length of the linac essentially is maintained and the energy spread is on the order of 10^{-4}. The energy spread is important for the generation of higher undulator harmonics The smaller the energy spread, the more intense are the higher undulator harmonics. Longer undulators also require a low emittance and low energy spread to reach their full potential. Calculated average brilliance curves for a low energy (3GeV) and a high energy (5GeV) ERL are shown in Fig. 22. (13)

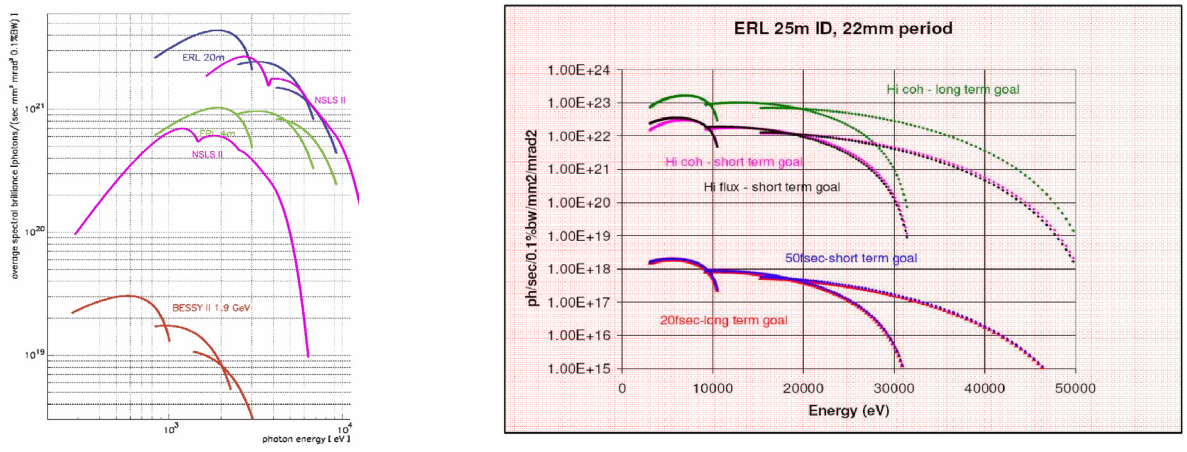


Fig. 22: *Calculated average brilliance for a 3 GeV ERL (left) and a 5 GeV ERL (right)(13)*

The average brilliance of these machines is expected to reach more than 10^{21} at 1 KeV and more than 10^{23} at 10 KeV photon energy. This is substantially better than any existing or planned storage ring facility (see Fig.14), including the ones under construction now. It almost reaches the limit of 'zero emittance' accelerator source as shown in Fig. 14. The improvement of the ERL compared with the storage rings is more significant at higher photon energies and also at higher electron beam energies, as expected. At low photon energies, corresponding to longer wavelengths, the photon source size is diffraction limited. This is the case whenever the condition $\varepsilon < \lambda/4\pi$ is met. Under these circumstances, the reduction of the emittance of the electron beam alone does not improve the brilliance of the source.

This however does not include all effects. In a long (25m) undulator, for example, **partial lasing** of the electron beam is to be expected for ultra bright electron beams. This is the same effect as in the FEL described above, but saturation is not reached. Nevertheless the brilliance can increase by several orders of magnitude. The prediction of the brilliance of ERL's in the XUV range accordingly requires detailed dynamical modeling of the electron beam and therefore this is beyond the scope of this brief review**.**

Seeding the photon beam of such a long undulator with an external laser should also be considered an option, which offers better control over the photon pulse shape and the synchronization to an external laser for pump-probe experiments.

It should be mentioned, that along the beam path around the arc of the ERL, the beam quality will deteriorate, depending on the extraction of photons and the insertion devices installed. The electron bunch properties are not quite the same at the end of the ARC as immediately after the linac. Care has to be taken, that this does not prevent the energy recovery process.

Another advantage of the ERL vs. the storage ring is the **round cross section of the beam** and the photon source. This is a specific advantage for any imaging application. Furthermore, there is no need to have a large horizontal aperture, which is required for the injection into the storage ring. Accordingly different types of undulators, for example the APPLE III type, may be constructed and implemented into an ERL. This undulator type has a substantially higher field strength on axis.

These are not the only advantages of the ERL design compared to storage rings. In general, the ERL offers also a great **flexibility** as far as the bunch pattern, repetition rate and especially **short pulse** characteristics are concerned. According to present designs ERL's will operate at an RF frequency of 1.3 GHz. This also is the maximum pulse repetition rate. The ERL offers the possibility to make use of bunch compression techniques to substantially shorten the electron bunches to produce photon **pulses as short as 10 fs** (at a reduced charge). In general the ERL offers a great flexibility as far as the electron optics and bunch shaping are concerned. This is inherent in the ERL design, since the electron beam parameters of the electron gun and the linac are essentially the ones determining the photon source. In the storage ring the photon source is given by the equilibrium reached after thousands of turns. Thus the electron beam in the storage ring has lost any memory of its source properties.

As far as **polarization**, **wavelength tuning**, and **resolution** of the photon source are concerned, the ERL has properties quite similar to a storage ring source. ERLs can satisfy the needs of a large user community *and* enable fascinating new science not possible with the existing machines. The ERL is especially suited for experiments requiring high average brilliance of the photon beam, such as microscopy, microspectroscopy, or high resolution electron spectroscopy. Compared to an FEL which is able to reach similar values in average brilliance, the photons of the ERL are spread over at least 10^4 times the number of pulses. Thus the danger of non-linear effects in the sample or in the detector is greatly reduced. Over large wavelength ranges the ERL achieves performance values close to the theoretical values of the ultimate, zero emittance accelerator based sources.

4. Critical Discussion: Ultimate storage rings, ERL, FEL's; where is the difference?

As stated above, the performance of storage rings is ultimately limited by the equilibrium properties of the stored beam. The electron beam emittance scales with 1/R of the bending radius, but this also imposes a physical size limit. Thus the quest for ultimate performance leaves the linac driven light sources as the sources of the 4th generation, to replace and complement todays storage rings. **The FEL's offer unique science opportunities, whenever a high peak brilliance is crucial.** This is for example the case for the study of non-linear phenomena, ultra low density dilute samples, and whenever the short fs-pulse characteristics are required. Seeding the FEL beam with an external laser driven source will substantially enhance the quality and reproducibility of the pulse shape of the FEL pulses. This also will improve the synchronization for pump probe studies. **Whenever a high average brilliance is the key requirement for the investigations, then the ERL is the source of choice.** These experiments include many imaging applications as well as high resolution electron spectroscopy, where space charge effects are a concern. For time resolved investigations the ERL offers 10 fs pulses, whenever the peak brilliance is not required. Schematically the accelerator based light sources are compared with storage rings schematically in Fig. 23

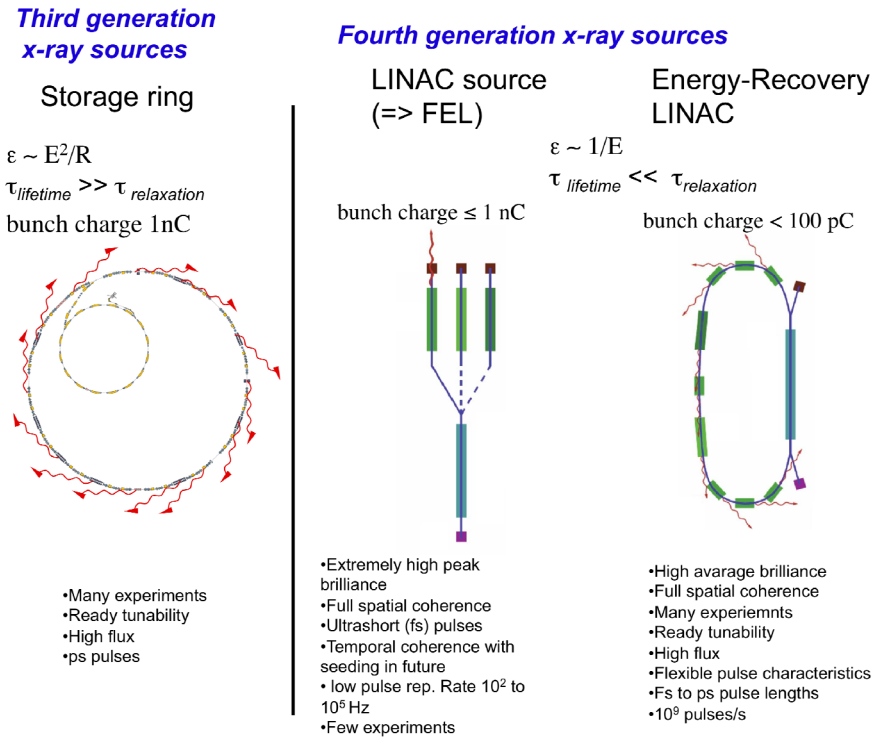


Fig. 23: *Schematic comparison between storage ring based and accelerator based X-ray light sources*

The ERL offers a high flexibility. It can be operated with multiple guns and in quite a set of different modes, emphasizing either high average brilliance, high coherence, or short pulses. Furthermore, pulses can be extracted from the ERL and run into long FEL undulators. Even if the peak current of this accelerator does not match the extreme conditions of the X-ray FEL,

lasing is achieved by setting up an optical cavity for the X-ray pulses with a clever design of back reflecting crystals. Such a scheme has been proposed and is outlined in Fig. 24

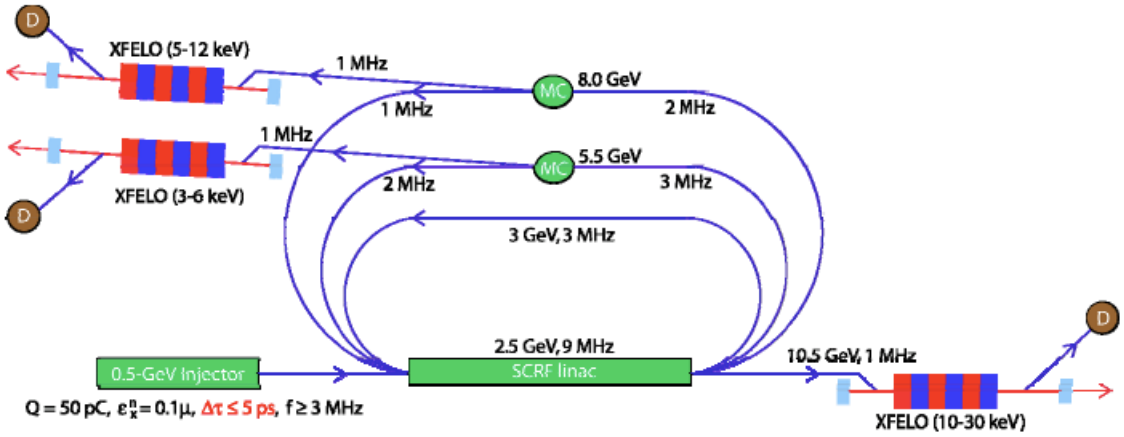


Fig. 24: *Scheme of combining a superconducting linear accelerator with several return arcs and various X-FEL oscillators (14)*

This scheme is included here to demonstrate the ultimate flexibility of using a superconducting linear accelerator as the basis for a future accelerator driven X-Ray light source facility. There are many clever ideas and concepts around and more will be generated. One just has to realize that the source of the photons in all these schemes is an electron bunch, shortened to sub ps bunch length with a very low emittance, and high peak current. Once the charge of this bunch is modulated at a certain wavelength, it turns into a laser medium at that wavelength and all the electrons start to emit coherently.

This points to an ideal tool kit for assembling an X-ray source in the future:
(1) a superconducting linear accelerator
(2) various electron guns for injection with different repetition rates, bunch charges, and bunch compression schemes
(3) one or more return arcs equipped with a series of sophisticated undulators for multiple experiments, and
(4) several fast kickers to extract single bunches into long FEL undulators.

This is a tool kit and naturally the tools chosen for any particular realization will not encompass all these elements but have to be chosen to match the science target to be addressed.

Why does such a source not exist yet? While an ERL for the IR range has been successfully operating at Jefferson-Lab and at the Budker institute in Novosibirsk for quite some time, there exists no X-ray facility based on an ERL. Simply, recirculation and energy recovery has not been demonstrated at the level of performance required for the X-ray ERL facilities. This leads to the question why recirculation is required at all ---- why not dump the beam as it is done in the FEL. At an average beam current of 100 mA and an energy of 5 GeV the total energy of the electron beam is 500 MW. Without recovering almost all of this energy this would mean that the total RF power installed would have to be in the GW range. Recirculation of the electrons to regain the energy is a must.

The FEL bunch cannot be recirculated. After saturated lasing has been achieved the brilliance of the beam is substantially reduced and the energy spread is also enlarged. On the other hand

for FEL's dumping the beam is not such a problem. The LCLS beam (120 pulses/s at 15 GeV, 3KA peak current and 200 fs length) has a total stored energy of about 1 KW. Even the European X-FEL, which aims at up to 10^5 pulses/s has a total stored energy of about 1 MW.

In addition to the demonstration of the recovery of the energy on a quantitative scale, the development of an X-ray ERL as well as the realization of FEL designs with MHz pulse rates presently is also hindered by the lack of technical developments as far as gun performance is concerned. Today's guns cannot yet deliver the bunch charges, emittances and repetition rates required for the full ERL or FEL design outlined above.

Both these technology goals, on the other hand, are quite well defined and within range. Once these have been achieved there will be a very bright future for accelerator based X-ray sources.

References

Section 3 and 4 were largely taken from ref. 6, the DOE report 'Next Generation Photon Sources for Grand Challenges in Science and Energy' DOE-BESAC (May 2009), where this text was contributed by the author

(1) X-Ray Data Booklet, Center for X-ray Optics and Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley USA 2009 (http://xdb.lbl.gov/xdb-new.pdf) and references therein
(2) WIKIPEDIA and J.D. Jackson, Classical Electrodynamics (1999)
(3) H. Winick, A. Bienenstock, Ann. Rev. Nucl. Part. Sci. **28**, 33 (1978)
(4) For a review of wigglers and undulators see: G. Brown, K. Halbach, J. Harris, H. Winick, Nucl. Instrum and Methods **208**, 65 (1983)
(5) J. Bahrdt, W. Frentrup, A. Gaupp, M. Scheer, W. Gudat, G. Ingold, S. Sasaki, Nucl. Instrum and Methods **A467**, 21 (2001)
(6) DOE; Next Generation Photon Sources for Challenges in Science and Energy DOE-BESAC (May 2009) (http://www.sc.doe.gov/bes/reports/list.html)
(7) R.W. Schoenlein, S. Chattopadhyay, H.H.W. Chong, T.E. Clover, P.A. Heiman, C.V. Shank, A.A. Zholents, M.S. Zolotorev Science **287**, 2237 (2000)
(8) K. Holldack, S. Khan, R. Mitzner, T. Quast Phys. Rev. Lett **96**, 054801 (2006)
(9) The Technical Design Report of the European XFEL (http://xfel.desy.de/tdr/tdr)
(10) L.H. YU, J. Wu , Nucl. Instrum and Meth. A 483, 493 (2002)
(11) The BESSY Soft X-Ray Free Electron Laser, Technical Design Report (BESSY, Berlin March 2004, ISBN 3-9809534-0-8)
(12) R. Mitzner, B. Siemer, M. Neeb, F. Siewert, S. Rolling, M. Rutkowski, A.A. Sorokin, M. Richter, P. Jurancic, K. Tiedtke, J. Feldhaus, W. Eberhardt, H. Zacharias, Optics Express, **16**, 019909 (2008)
(13) G.H. Hoffstaetter, V. Litvinenko, H. Owen, Nucl. Instrum and Methods **A557**, 345 (2006)
(14) W.A. Barletta et al., Nucl. Instrum. and Methods, **A618**, 69 (2010)

C 3 Instruments for Neutron Scattering

Michael Monkenbusch
Jülich Center for Neutron Science
Forschungszentrum Jülich, D-52425 Jülich

Contents

[0]Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

In general [1] a neutron scattering instrument measures the double differential cross section

$$\frac{d^2\sigma}{d\Omega dE'} \;=\; N\,\frac{k'}{k} b^2 S(\mathbf{Q},\omega) \tag{1}$$

where k', k denote the wavevector lengths of the scattered (k') and the incoming neutrons. For a simple monoatomic bulk sample N is the number atoms and b their scattering length. The available physical information may be expressed in terms of the scattering function $S(\mathbf{Q},\omega)$. The energy transfer that occurs during the scattering process $\hbar\omega = (E - E')$ is given by the difference of the energy E of incoming and scattered neutrons (E'). The momentum transfer $\hbar Q$ is proportional to the difference $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$ of initial and final wavevectors. Also the energy transfer is related to the change in these wavevectors

$$\hbar\omega = (\hbar^2/2m_n)(k^2 - k'^2) \tag{2}$$

according to the relation between $\mathbf{k}$ and the neutron velocity $\mathbf{v}$ and neutron wavelength λ

$$\mathbf{v}m_n \;=\; \hbar\mathbf{k} \quad \text{and} \quad \frac{2\pi}{\lambda} \;=\; k \tag{3}$$

The energy transfer may be obtained either by direct neutron velocity measurements using their flight times or –using the wave properties– by analysis of the wavelength by e.g. Bragg reflection from a periodic structure (crystal, multilayer). Instruments as diffractometers that focus on the microscopic structure of the sample mainly analyze the $\mathbf{Q}$ dependence of the scattering without energy analysis. The same is true for reflectometers. On the other hand spectrometers focus on the energy transfer $\hbar\omega$ during scattering and its dependence on the momentum transfer $\mathbf{Q}$.

Depending on the application and the abilities of the used techniques as well $\hbar\omega$ as $\mathbf{Q}$ are more or less precisely defined. The different combination of *resolutions* with respect to $\mathbf{Q}$ and ω yields the variety of instruments, in particular since the realized resolutions determine the maximum available intensity. Neutron scattering always is an intensity limited method, even for the most powerful research reactors or spallation sources. Thermal or cold neutrons emerge from a moderator which exposes an extended area (some $100\,\mathrm{cm}^2$) emitting neutrons in all direction and with a velocity distribution according to the effective moderator temperature. Therefore a good adaption of resolution of an instrument to the requirements of the physical question such that all acceptable neutrons from the source are utilized is a key to perform an experiment with optimum intensity on the detector.

Neutrons are spin 1/2 particles with an associated magnetic moment. The scattering cross sections of a sample may depend on the spin state of the incoming neutrons and the scattering may change the spin state. In particular samples with magnetic structures will show these kinds of polarization dependent scattering and analysis of it can reveal details on magnetic structures and excitations.

[1] Note that the majority of neutron instruments at reactor or spallation sources are scattering instruments including reflectometers. However, there are also some other applications as radiographic and tomographic imaging or the study of nuclear reactions due to neutron capture for chemical analysis or nuclear physics experiments. Finally some fundamental physics apparatus are devoted to the properties of the neutron itself as e.g. precise measurement of it's lifetime or the search for an electrical dipole moment. These are not covered by the present article.

In the following of this chapter a selection the most important instrument types to investigate structures in condensed matter samples are presented. These *diffractometers*, which do not analyze the energy transfer but detect all neutrons that are scattered by the sample, are explained. The structure information is obtained from the scattering angle or more precise the **Q** dependence of the intensity. The implicit diffractometer assumption in general is that there is only negligible or no energy transfer.
Later *spectrometers* that focus on the analysis of the energy transfer during scattering are described. Here motion and motional patterns on atomic scale or excitations are addressed. Depending on the instrument the **Q** resolution is worse than those of a diffractometer in order to collect sufficient intensity to enable the spectral analysis. The covered energy transfers (not by a single instrument) range from 10% down to 0.1% of the initial neutron energy for instruments with direct spectral analysis, e.g. by Bragg reflection, and further down into the ppm regime for neutron-spin-echo spectrometers.
First the most important functional components that are part of neutron scattering instruments are now presented.

2 Components

2.1 Divergence Control

Neutron sources with sufficient intensity for scattering instrumentation are either research reactors or spallation sources. Neutrons with suitable wavelength for scattering experiments on condensed matter are obtained by down-scattering and thermalization in a moderator (see chapter C1). The moderators emit a spectrum with broad energy spread of neutrons, which is close to a Maxwell distribution corresponding to the effective temperature plus a contribution of so-called epithermal neutrons that still are not fully slowed down. The neutrons emerge from an extended area (surface of the moderator). The moderator emits neutrons without preferred direction.
Instruments receive the neutrons through channels in the several meters thick radiation shielding leading to the moderators. They have typical cross sections of $\simeq 10 \cdots 20\,\mathrm{cm}$ and several m length. In order to obtain a suitably well defined wavevector, neutrons of a certain direction have to be selected. A process called collimation (see fig. 1). The simplest way would be to rely on entrance an exit apertures (“pin-holes”) in a certain distance. And in a way the channel in the shielding may serve as a coarse collimator. If however this it not sufficient other types of collimation, in particular Soller collimators [1] are used. These collimators consist of an array of parallel absorbing sheets in a frame. Only a few isotopes are available to realize effective absorbers from which pin-holes, diaphragms aperture blades and collimator sheets can be made. For thermal and cold neutrons these parts preferentially contain $^{10}\mathrm{B}$, Gd or Cd and eventually $^{6}\mathrm{Li}$. Here material thicknesses in the range from a tenth to a few mm usually suffice. A good collimation yields a narrow spread of neutron directions in the beam on the expense of intensity. If space around a neutron source has to be shared by many instruments or if for a pulsed source a certain distance is required for a time-of-flight tagging of the neutron velocity, a large distance between source (moderator) and sample often will result. In most of these cases the “natural” collimation that results from the ratio of moderator(opening) plus sample sizes compared to their distance is much narrower than the instrument requires. To avoid the intensity loss due to this mismatch of beam divergence so-called *neutron guides* are used to bridge the

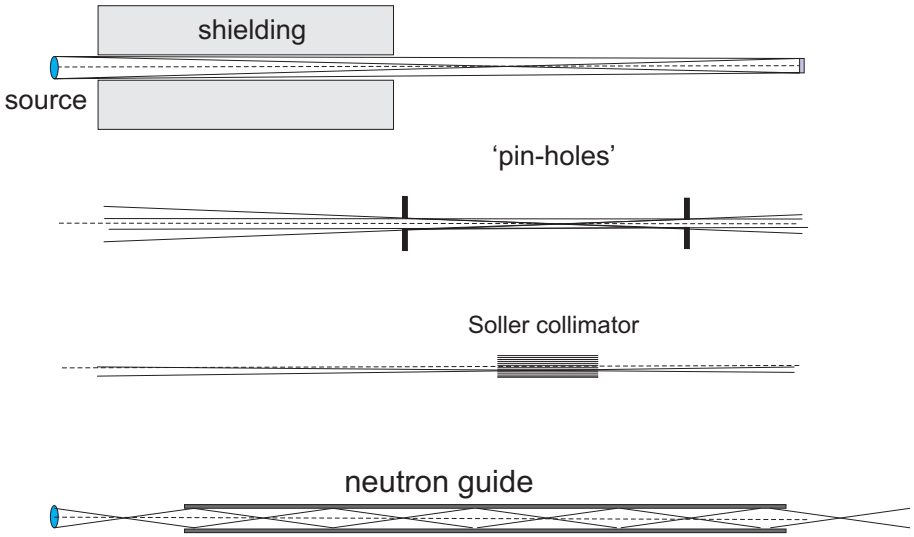


Fig. 1: *The divergence of a neutron beam is naturally limited by the effective source size or the beam channel width, the sample (target) size and their distance. Explicit control can be obtained by pin-holes, diaphragms or slits at a distance, in those cases the beam diameter and the degree of collimation determines the required length. Finally compact Soller collimators with many parallel (or radially converging) channels decouple length, beam width and achieved divergence. Finally neutron guides (channels with reflecting surface, see text) are used if the natural divergence would be to small.*

distance. Neutron guides consist of evacuated (rectangular) channels with very smooth coated inner surfaces. The coating is such that neutrons that hit the surface at a sufficiently flat angle are reflected. For most elements the effective index of refraction, $n = 1 - \lambda^2\, b\, N/(2\pi)$ (with b=scattering length and N=number density of nuclei) for thermal neutrons is less than 1. This means that below a limit angle of incidence neutrons undergo total reflection at a polished surface of such an element. The standard coating that utilizes total reflection is a layer of several 100nm of Ni, the angle of total reflection may be increased by 20% if $^{58}\mathrm{Ni}$ is used instead of natural Ni. Even larger reflection angles are realized by multilayer coatings as Ni/Ti. The ratio of the maximum reflection angle of a given (multi)layer and a standard Ni coating usually is denoted by $\mathbf{m}$. Guides with **m**-values between 2 and 3 are readily available 5-6 seems technical feasible. However, the required number of layers, i.e. the effort and cost increases steeply with m. For a given layer the maximum reflection angle is proportional to the neutron wavelength and the value is about 1^0 for $10\,\text{Å}$ neutrons for a Ni coating.

2.2 Spectrum Control

Besides the direction selection a well defined initial wavevector in addition requires the selection of a specific neutron wavelength, i.e. neutron velocity. The method employed in the first neutron instruments and which still is widely used for this purpose is the Bragg reflection by large (mosaic) crystals. Nowadays widely used are pyrolytic graphite, PG(002), Si(111,311), Cu(111,200,220,331), Ge(hhl) and for polarization Heusler alloy $\mathrm{Cu_2\,Mn\,Al}$(111) crystals. Since the primary neutrons have a continuous wavelength distribution the intensity of the reflected beam is proportional to the width of the selection band. A very perfect crystal

yields a well defined selection, however, with little intensity. Therefore for many applications in neutron scattering imperfect crystals with a "mosaic" of small slightly misaligned domains are preferred or even essential. Whereas for Si special effort and skill is needed to "deteriorate" the available nearly perfect crystals in a defined manner into a mosaic crystal, for pyrolytic graphite the manufacturing process rather yields crystals at the other margin of perfection. For most purposes where cold neutrons with about 1% wavelength definition are needed graphite is the first choice. For shorter wavelength neutrons e.g. germanium or copper crystals may be used. See figure 2.
Depending on application the selected wavelength λ may be contaminated by intensity from higher order reflection $\lambda/2,\ \lambda/3, \cdots$. In those cases additional filters are needed that are based on crystalline Bragg reflection (e.g. Be-filter [2]) or a mechanical velocity selector [3, 4].

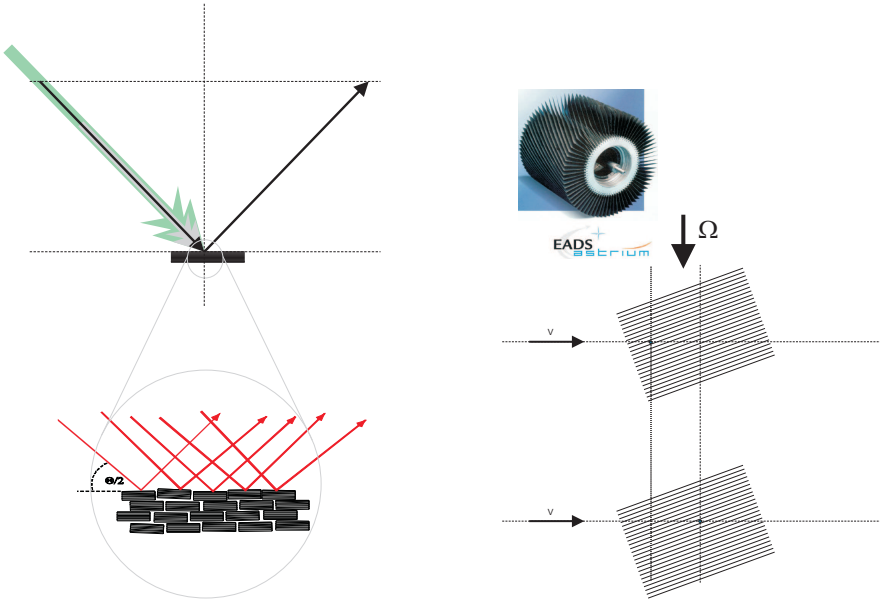


Fig. 2: *A suitable means to select a narrow wavelength respectively velocity band from a "white" beam of neutrons is: left:the Bragg reflection from a crystal. Suitable crystals are e.g. Graphite (002), Copper, Germanium etc.. Or right: a mechanical velocity selector. A inclined collimating structure, usually helically wound lamellae on the periphery of a rotating cylinder move perpendicular to the beam direction. A certain open channel section along the beam axis therefore moves with a velocity* $v = v_r/\tan\alpha = \Omega\, r/\tan\alpha$*, with* α *the inclination angle,* v_r *the velocity of the cylinder periphery at radius* r*.*

But not only the wave property of the neutron allows a selection of wavelength. Since the wavelength is proportional to the velocity and the typical neutron velocities range between several 100 m/s to a few 1000 m/s it may also be chosen by controlling the neutron time-of-flight over a distance of a few meters. This can be achieved by means of choppers that open the neutron beam path periodically for a short time. The delay between different choppers along a beam path thus determines the velocity of the transmitted neutron bunches. Disc choppers consist of a rotating absorbing disc with neutron transparent sectors and rotating axis parallel to the beam direction. So called Fermi-choppers [5] on the other hand are small rotating collimators that for a short period of their revolution are aligned with the beam direction. The rotation axis is perpendicular to the beam direction. Disc choppers need some time to switch a larger beam

cross section from closed to open whereas Fermi-choppers may do this in a shorter time. However, their collimation direction changes at a high rate and thus slow neutrons may be hit by an absorbing wall before they leave the chopper again. This effect may be utilized to select a limited velocity band by bending the collimation channels.
Even more this effect is used in mechanical velocity selectors (figure 2), where absorbing lamellae in a twisted configuration are on the periphery of a rotating cylinder. If the neutron velocity matches the apparent speed of the gap between two lamellae on a path parallel to the rotation axis, the neutron is transmitted without being hit by an absorbing wall. Neutrons with differing velocities are absorbed. Unlike a chopper the velocity selector performs a coarse monochromatization for a continuous beam. Typical monochromatizations are about 10% of the mean velocity.

2.3 Polarization

The first methods used to prepare and analyze polarized neutron beams rely on the magnetic part of Bragg reflections [6, 7] including the use of magnetized polycrystalline iron as polarizing filter. Today Heusler crystals of the type $\mathrm{Cu_2\,Mn\,Al}$ are used as efficient polarizing monochromators for thermal neutrons. Cold neutrons rather are polarized by utilizing the spin dependent reflection from magnetic multilayers, as illustrated in figure 3 . Common A-B multilayer systems are FeCo-Ti and Fe-Si [8], the choice is made such that the contrast between the magnetized A and the nonmagnetic B layer vanishes for one spin state and has a sizable value for the other. Only neutrons with spin states that “see” the contrasts are reflected by the multilayer stack. The layer thickness is varied over the stack such that reflection is large over a wide range of incident angles respectively wavelengths up to a maximum. The maximum value is specified in terms of multiples m of the total reflection edge of Ni. m values up to and even beyond $m = 4$ are state of the art.

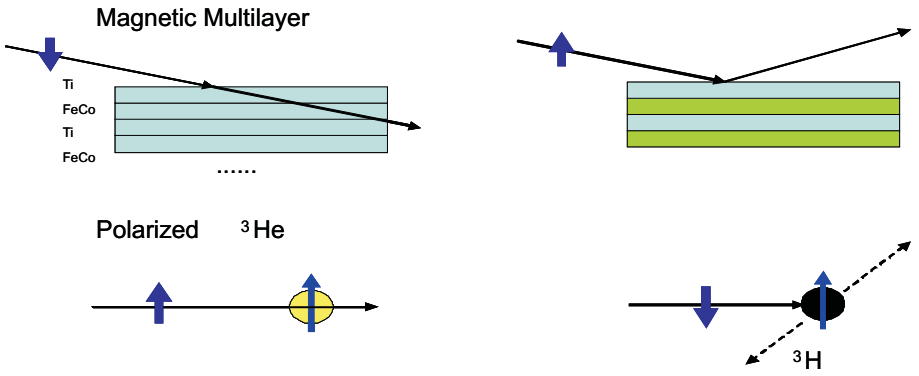


Fig. 3: *Illustration of two important techniques to polarize neutron beams. Upper part: layering of a magnetic and nonmagnetic material with layer distances in the several nm range is used to yield a reflecting structure for one spin component. The neutron spin interaction with the magnetization adds or subtracts to the nuclear potential and is adjusted such that for one spin state the layering becomes invisible. Bottom: an alternative way to polarize neutrons is to utilize the fact that the absorption of neutrons by ^{3}He is spin dependent. Here the challenge is to establish and preserve a high nuclear polarization of the ^{3}He nuclei.*

Finally neutrons may be polarized by absorption from nuclear spin polarized ^{3}He. The high

absorption cross section of ^{3}He holds only for one spin state. The absorption cross section is proportional to the neutron wavelength. Special quartz cells containing spin-polarized ^{3}He at pressures close to 1 bar can be position in the neutron path as polarization filters. The advantage of a polarized ^{3}He filter is its wide angular acceptance even for short wavelengths. The technical difficulty consists in creating and maintaining the polarization of ^{3}He nuclei either off-line (MEOP) [9] or on-line in a steady state setup (SEOP) [10]. Preservation of ^{3}He polarization requires a very homogeneous magnetic field and selected cells with low magnetic impurities exposed by the inner wall. Since the maximum nuclear polarization is significantly less than 100% and in off-line cells decays with time constants of a few 100h the thickness and pressure of the gas volume has to be optimized for the wavelength range to be used. Many of the various scattering techniques that are presented in the following can be used also in combination with polarization analysis.

2.4 Neutron Detection

Efficient neutron detection is as important as the production of a high flux and efficient transport to the sample. The goal here is to convert close to 100% of those neutrons that reach the detector into counts. In addition as much of the available solid angle around the sample should be covered with detectors. In order to be able to extract physical information from the detected neutrons the detector must yield information on the position of the detection event as well as on the time. The latter is needed for all time-of-flight based methods, which are particularly important at pulsed (spallation) sources.

The kinetic energy of a thermal neutron is not sufficient to create a detectable signal in a counting tube or scintillator. Therefore detectors for neutron scattering rely on the energy release of a nuclear reaction, namely the absorption of the neutron. A suitable detector material must have a large absorption and the energy release should be such that it causes large ionization in a counting gas or intense light output in a scintillator. Only a few isotopes, listed in table 2.4, proved suitable for this purpose.

The use of a single detector that scans the scattering angle range step by step is a quite inefficient collection method and is replaced by the use of multichannel detectors that cover a whole range at once. The techniques used for this purpose range from assemblies of many single counting tubes, integrated counting wires in a sector covering gas detector (“banana”) to scintillation counters with position sensitivity. It was a long way from the first single BF_3 tube neutron counters of the early diffractometers and 3-axis spectrometers to modern large area position sensitive detector arrays. This development was fostered by the huge progress electronics, data acquisition and handling made in-between. The GBytes from square meters of position-sensitive and time resolved detectors are just becoming treatable. The use of ^{3}He as detection gas that serves as neutron converter and counting gas with low γ-sensitivity also was a big progress. Due to a severe shortage of ^{3}He which recently became obvious, now new concepts for efficient neutron detection are being investigated and developed in order to preserve the performance of neutron scattering instrumentation. Besides gas counters scintillation detectors with ^{6}Li as neutron converter in scintillating glass or mixed with ZnS as scintillator are used. Either they are combined with photomultiplier arrays or CCD cameras. The latter without single event counting and time resolution.

Image plates containing a neutron converter (Li or Gd) store the energy from neutron absorption events in a phosphor which after exposure (similar to a photographic film) is read-out by stimulating light emission from the phosphor with a laser and detection of the emitted light.

Isotope	$\sigma_a/10^{-24}\mathrm{cm}^2$	reaction	energy release /MeV
^{3}He	5333	$\mathrm{n} +^3 \mathrm{He} \rightarrow \mathrm{p} + \mathrm{t}$	0.76
^{6}Li	940	$\mathrm{n} +^6 \mathrm{Li} \rightarrow \alpha + \mathrm{t}$	4.8
^{10}B	3837	$\mathrm{n} +^{10} \mathrm{B} \rightarrow \alpha +^7 \mathrm{Li}^*$	2.3
^{155}Gd	60900	$\mathrm{n} +^{155} \mathrm{Gd} \rightarrow^{156} \mathrm{Gd} + \gamma\mathrm{s}$	conversion e $\simeq$ 60 KeV [11]
^{157}Gd	254000	$\mathrm{n} +^{157} \mathrm{Gd} \rightarrow^{158} \mathrm{Gd} + \gamma\mathrm{s}$	"
^{235}U	680	fission	170

Table 1: *Reactions that are used to detect thermal neutrons. Absorption cross sections are given for neutrons with v=2200m/s.*

3 Diffractometers

Neutrons are especially useful if hydrogen or other light element positions or magnetic structures are to be determined. Combination with X-ray diffraction then enables the separate visualization of atomic position and electron density maps.

3.1 2-axis diffractometer, powder diffractometer

In its essence the generic 2-axis type neutron diffractometer consists of a collimator that defines the beam direction and a (mosaic) crystal oriented such that Bragg reflection for neutrons of desired selected wavelength occurs into the collimated direction. The monochromatized neutrons then hit the sample and are scattered in different directions. A crystalline powder sample would produce intensity on Debye-Scherrer rings. The early simple neutron diffratometers used a single detector e.g. a BF_3 counting tube that scan the scattering angle, around the sample. Count rate detection as a function of scattering angle 2θ yields the diffraction pattern. Modern instruments [12, 13, 14] rather have position sensitive detector (arrays) that cover a large range of solid angle and different scattering angles. The corresponding lattice plane distance follows from

$$d = \frac{\lambda}{2\,\sin\theta} \tag{4}$$

The resolution is given by

$$\frac{\Delta d}{d} = \sqrt{\left(\frac{\Delta\lambda}{\lambda}\right)^2 + \left(\frac{\Delta\theta}{\tan\theta}\right)^2} \tag{5}$$

High resolution requires a well defined incoming wavelength and a well known scattering angle –except close to $2\theta = \pi$ –. In general one may scan the d-value also by changing the wavelength λ instead of scanning the scattering angle 2θ. At a pulsed source the λ tagging is given by the difference between start time of the pulse and neutron arrival time at the detector, $\lambda(\Delta t) = \Delta t\, h/(m_n\, L)$. h is the Planck constant, m_n the neutron mass and L the distance from pulse source and detector. Combination of an (area) detector array and the time-of-flight method yields a high detection data rate and is commonly used at pulse spallation sources [14, 15]. Even at a continuous reactor source the loss associated with a pulse chopper may pay-off [16]. At a short pulse spallation source very high resolution can be achieved by using a long flight path ($\simeq$100m) to tag the wavelength and concentrating detectors in the back scattering region,

i.e. $1/\tan\theta \to 0$. Examples for diffraction results obtained at the instrument POWGEN at the SNS is shown in figure 4.

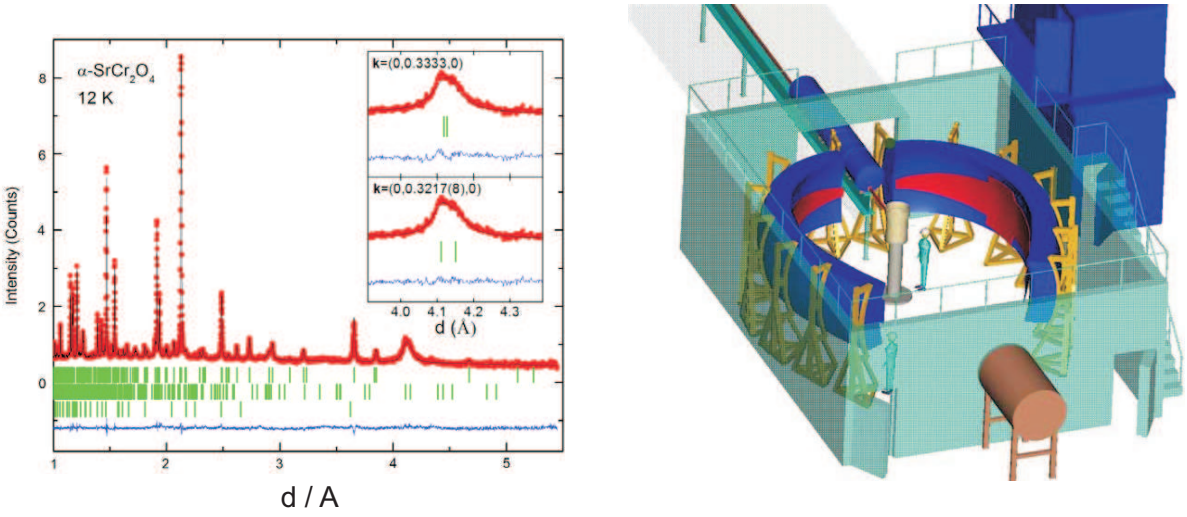


Fig. 4: *Example for a powder diffraction pattern as obtained at then time-of-flight diffractometer POWGEN at the SNS in Oak Ridge. The low temperature pattern here exhibits occurrence of magnetic reflections of an incommensurate helical structure around* $\mathrm{d} = 4.1\,\text{Å}$ *that are absent above the magnetic ordering temperature of 43 K. Precise determination of the lattice constants and correlating with the magnetic features indicate the presence of a magneto-structural coupling. [17] At the right side a sketch of the instrument is shown. The white neutron beam starts in the moderator at 60 m distance from the sample, it arrives through a neutron guide at the sample located in the center of the detector arrangement. The curved surface indicates the detector area, which is pixelated into about 240000 cells. Software correlates arrival time and pixel position and is used to accumulated data from a larger detector surface. The final result is seen in the diffractogram at the top.* **Figure (left part) from ref. [17] with permission, copyright IOP.**

3.2 Single crystal diffractometer

Single crystal diffractometers are used to measure the intensity of as many Bragg reflections in reciprocal space as possible in order to determine the structure of the unit cell. For crystals with small unit cells Bragg reflection intensity is only detected if both the scattering angle between incoming and detected neutron beam has the right value and in addition the crystal orientation must be such that the corresponding lattice planes fulfill the reflection condition. Therefore these spectrometers need a goniometer for sample orientation in addition to the ability to set the scattering angle.

Giving the sample orientation and the positioning of an arc of detectors more freedom to rotate around tilted axis and go out of the horizontal plane makes it possible to map the intensity of a selected plane in reciprocal space. This so-called *flat-cone* instrument [18] in particular also enables easy access to scattering intensity in-between the Bragg reflections in a reciprocal lattice plane (u,v) as e.g. $(h \neq 0, u, v)$.

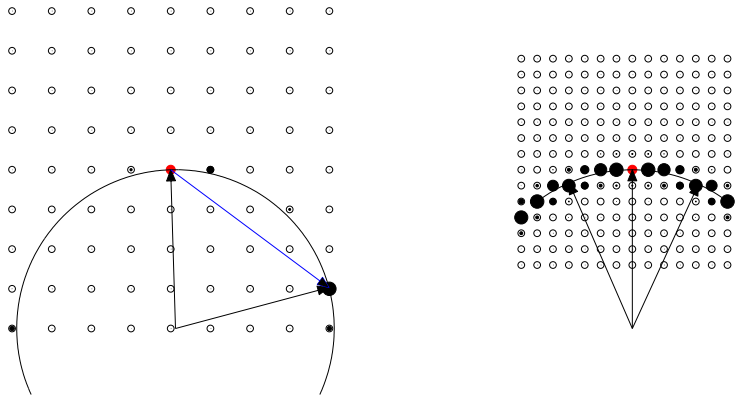


Fig. 5: *Ewald spheres in a reciprocal lattice from a small unit cell and from a large unit cell crystal (e.g. protein).*

In the case of very large unit cells as for protein crystals the curvature of the Ewald sphere is low enough that many reflections fulfill the Bragg condition within the resolution width simultaneously. Here a position sensitive detection with a sub-mm resolution is needed for an efficient data collection. At a continuous source the detector can either be an image plate or a scintillator screen with CCD camera. Wound into a cylinder with the sample position in the center an image plate detector can cover a very large solid angle at once. In this geometry it is possible to read-out the image plate in place. With CCD or other planar detectors a comparable solid angle coverage requires the use of many modules that are positioned around the sample. Figure 6 show a drawing of the BIODIFF experiment at the FRMII in Garching that utilizes this detection method, a sample diffraction image is also shown.
The sample crystal is then rotated around one axis and data are taken for each rotation step. Eventually the sample crystal has to be remounted in another orientation to complete the data collection. Reflections then are automatically indexed by software. A protein crystal analysis requires the collection of several 10000 individual reflex intensities. Only the use of area detectors like image plate CCD or other high resolution detectors makes this type of data collection possible in periods of a few weeks [19].
At spallation or other pulsed sources neutrons from an extended range of the continuous spectrum are to be used ("Laue method"). Unlike for the "Laue" at a continuous source overlap of reflection and λ-tagging can be done by the neutron arrival time. This rules out image plate and CCD-detectors, which have no appropriate time resolution. For those instruments only highly sophisticated and rather expensive high resolution gas detectors or position resolving scintillation detectors are suitable [20, 21].

4 Small angle scattering diffractometer and Reflectometers

Small angle scattering is a diffraction method for the observation of mesoscopic structures in the range between 1 nm and several 100 nm. In particular the possibility to vary the scattering contrast by exchange of H with D makes it a unique tool for soft-matter investigations (polymers, bio-molecular systems, complex fluids, gels etc.) [22, 23]. For the spatial resolution of

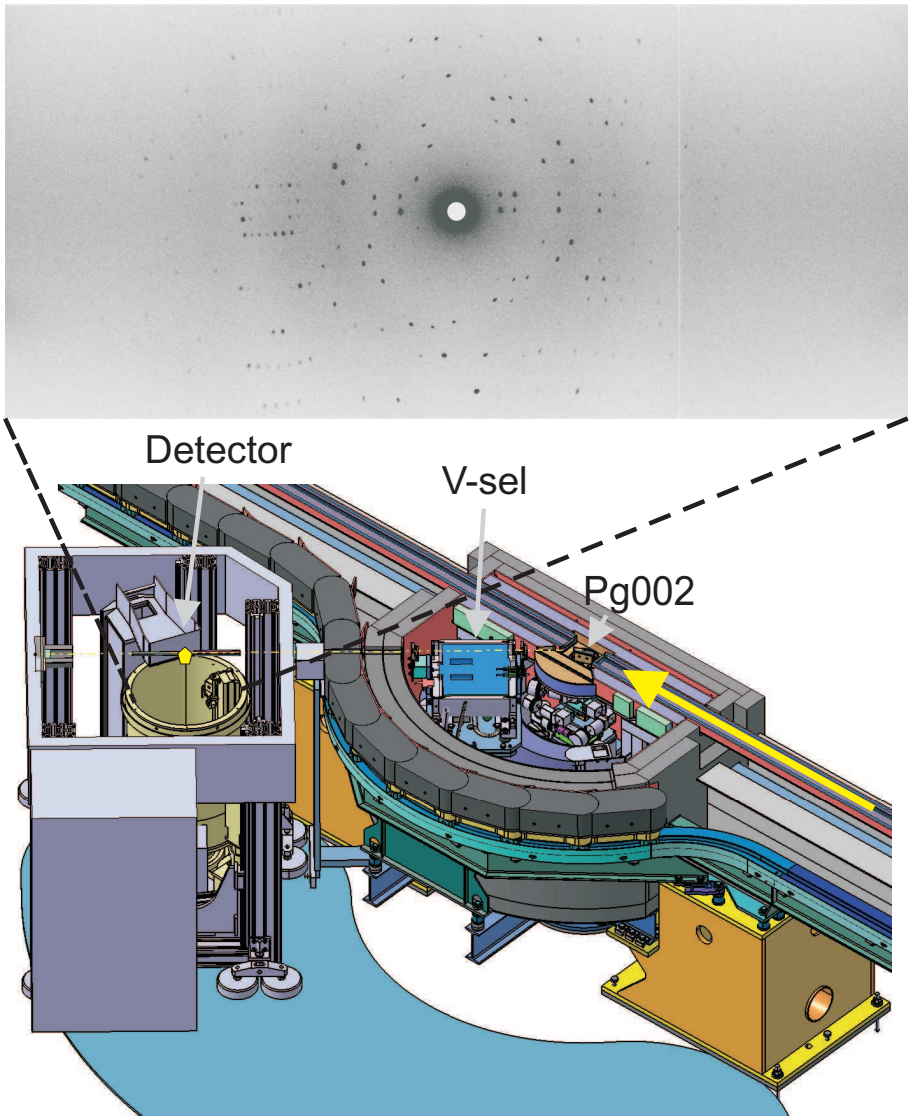


Fig. 6: *Reflections from a myoglobin crystal on the cylindrical image-plate of the BIODIFF macromolecular single crystal diffractometer at the FRMII in Garching. A considerable number of Bragg spots are seen for one sample orientation. A complete experiment comprises the data collection for many sample rotation angles and orientations in order to measure several 10000 intensities. The lower part of the figure displays a cut through the instrument. Neutron in the guide (indicated by the yellow arrow) are monochromatized and reflected by a graphite crystal, PG002. Higher order wavelength contaminations are removed by a velocity selector, v-sel. The narrow beam of a few mm diameter then leaves the lead shielding and enters the detector rack. The yellow polygon indicates the sample crystal. In operation the cylindrical image plate detector is raised such that the sample is in its center. The configuration shown is for exposure of the auxiliary CCD camera.*

SANS the sample can be described by a scattering length density $\rho(\mathbf{r})$ or better by the contrast $\Delta\rho(\mathbf{r}) - \rho_0$. The scattering length density results form assigning the sum of all nuclear scattering lengths, b_l within a molecular building block to the volume occupied by this block. ρ_0 is that of an embedding medium, e.g. a solvent. The the differential scattering cross section reads:

$$\frac{d\sigma}{d\Omega}(\mathbf{Q}) = \sum_{l,j}^{N} b_l b_j e^{-i\mathbf{Q}(\mathbf{r}_l - \mathbf{r}_j)} \quad \simeq \quad \left| \int \Delta\rho(\mathbf{r}) e^{-\mathbf{Q}\cdot\mathbf{r}} d^3\mathbf{r} \right|^2 \tag{6}$$

For colloidal objects in dilute concentration, as in the example in figure 8, the integral in eq. 6 can be done for one the colloidal particle objects multiplied by their number. For higher concentrations interference between objects modifies the cross section.

The principle of a *classical pin-hole camera* type small angle neutron scattering (SANS) instrument is rather simple. In its essence the proper camera consist of two pin-holes (entrance aperture and exit aperture) at a distance of several meters. The apertures have typical sizes of 3cm and 1cm and serve as collimator to define the direction of the incident neutrons. After hitting the sample closely behind the exit aperture, the scattered neutrons are detected by a position sensitive detector in a distance of 1 m up to several 10 m from the sample. The effective distance between the pin-holes and between sample and detector are variable in order to cover a broader Q-range. The basic setup is show in figure 7. Since neutrons are significantly scattered by air, it is necessary to evacuate the long flight spaces in the collimator and between sample and detector. Especially the latter implies that the appearance of a SANS instrument is dominated by a huge vacuum tank which contains the detector on rails to vary the distance for different Q-ranges and resolutions. The collimation length, i.e. the distance between exit and entrance aperture has to be chosen accordingly. This is done by introducing neutron guides into the space between velocity selector and current position of the entrance aperture. Thereby a beam with full divergence is lead until the effective start of the pin-hole collimation. The wavelength selection at a continuous source is performed with a typical relative width of 10% by the use of a mechanical velocity selector (see. 2.2). Typically a wavelength range between 5Å and 20Å is used for experiments. Example data from polystyrene spheres with different radii in dilute colloidal solution are shown in figure 8.

In principle a SANS experiment at a *pulsed source* looks quite similar to the one described above [24]. The main difference is that the velocity selector is omitted and replaced by the time-of-flight analysis of the neutrons arriving at the detector. The effective wavevector $\mathbf{Q}$ at a given detector pixel now depends on the arrival time of the neutron that tags its velocity. To arrive at the standard representation of SANS data in terms of $d\sigma/d\Omega(Q)$ grouping of pixels and time-bins yielding equal Q has to be performed by software.

To achieve instruments with *Ultra high resolution, USANS* one has to observe that the pin-hole collimation at the longest collimation distances yields a lowest $Q \simeq 10^{-3}\,\text{Å}^{-1}$. Since the size of both apertures enter this resolution limit and since the intensity is proportional to the product of both aperture areas, an increase of Q-resolution by this means is accompanied by a significant intensity reduction with the 4-th power of the resolution improvement. This effect poses a practical limit to the resolution. [2] Within a pin-hole instrument the use of neutron lenses can extend the Q-resolution. A focusing lens close to the sample position is used to

[2] A way out would be to make the instrument even longer. But the available space and the increasing costs set a limit. Note that the fact that the existing instruments are already very large are caused by this relation. Since the moderator is an extended source with a given brilliance intensity can only be gained by increasing the areas of source and sample. To keep the divergence in the limits posed by the resolution then requires large distances.

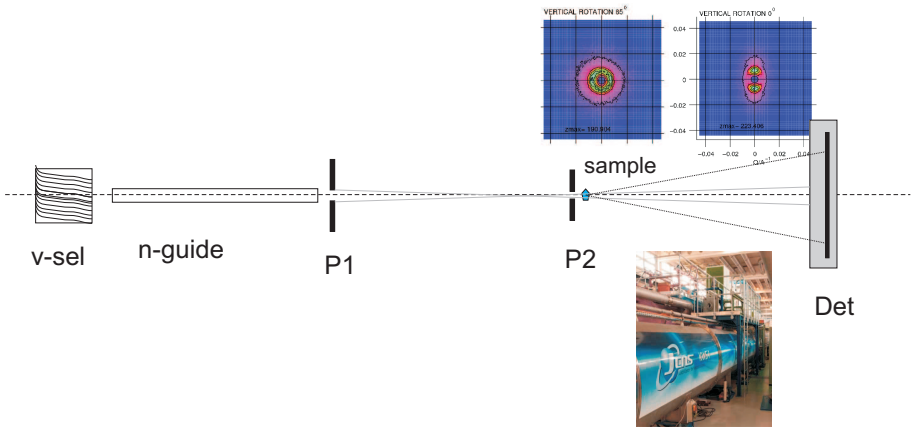


Fig. 7: *Schematics of a classical small angle neutron camera with pin-hole collimation. The collimation distance between the two pin-holes, $L_{\rm coll}$ is the same as the sample (at pin-hole P2) detector distance $L_{\rm det}$. After a velocity selector at the beginning a section with neutron guides bridges the distance to the first pin-hole. To vary the distance more or less neutron guide sections may be moved into the beam path. The detector can move in a vacuum tank to always match $L_{\rm det} = L_{\rm coll}$. Typical dimensions are $1 \cdots 40$ m for $L_{\rm coll}$ and 1cm for the sample pin-hole P1 and about 3cm for P2.*

produce an image of a small entrance aperture at the start of the collimation on the detector plane. The intensity loss due to the narrowing of the entrance aperture is mitigated by the larger aperture at the sample position which is only limited by either the lens diameter (some cm) or the available sample area or the divergence of the illumination of the input pin-hole. A neutron lens can be made from materials with low neutron absorption and large scattering length density $b\,N$. The small angle scattering of the used material and its surfaces must be very low. Single crystals of MgF are a viable choice. The refractive index $n = 1 - \lambda^2\, b\, N/(2\pi)$ is slightly smaller than 1. To get a reasonable focusing effect a larger number of biconvex lenses have to be stacked. In order to reduce intensity losses due to thermal diffuse scattering the lenses should be cooled [25]. By this means the lowest available Q can be extended by one order of magnitude. Limitations occur due to chromatic errors (i.e. the wavelength dependence of the refractive index and gravitation). In a *Focusing mirror SANS*, as illustrated in figure 9, chromatic errors can be avoided by using elliptical mirrors instead of lenses. The challenge here is the necessary quality of the surface such that its parasitic scattering stays orders of magnitude below the scattering intensity expected from a sample. [26] Other techniques as the *Double Crystal Diffractometer, DKD* or the *Spin Echo Small Angle Scattering, SESANS* even enable the exploration of scattering at very low angles that lead to real space dimensions in the μm range, which overlap with light scattering resolution. Due to the limited space we mention them only briefly here. First there is the double crystal method (DKD) which utilizes the reflection sequence from two perfect (silicon) crystals, which only happens when these are rotated to the same angle within μ rad. Minute deviations of neutron paths due to scattering lead to deviations of the rotation angle that transmits these neutrons. To obtain a diffraction pattern the crystal angle has to be scanned stepwise [27]. Wide range SANS data from a hierarchically structured sample obtained by a combination of pin-hole, focusing mirror and DKD SANS, are show in figure 10. Polymer containing wax crystallites arrange to larger aggregates that finally are even visible in a light microscope. Length scales from several micrometers down to the nm

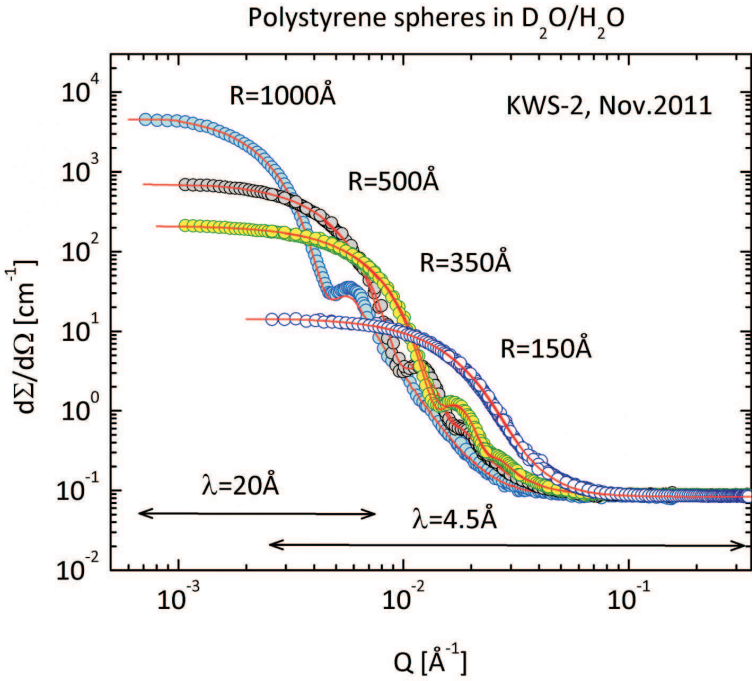


Fig. 8: *Reference examples of radially averaged SANS patterns from colloidal suspensions of polystyrene spheres of different sizes exploring the range of the typical pin-hole SANS-camera, here KWS-2 at the FRMII using different wavelength and collimation lengths.*

thick wax platelet size are covered. A different high resolution method utilizes the coding of small angular deviations in spin precession of the neutron (SESANS) [28]. As Fourier method it may be imagined similar to the projection of a more or less narrow sine stripe pattern on an integrating detector with a matched stripe mask. If small angle scattering distributes neutron such that their deviation is large enough to go from a maximum of a stripe to the adjacent minimum, the contrast (variation) upon shift of the stripe mask is reduced or vanished.

The specular reflection from a planar surface contains information on the scattering length density profile perpendicular to the surface, which is made accessible by *neutron reflectometers*. Beyond the total reflection, which depends on the scattering length density of the substrate, the Fresnel reflectivity $\propto Q^{-4}$ is modulated according to the Fourier transform of the density profile across the surface. Lateral structures in addition give rise to off-specular scattering (see also 4). In order to measure the specular intensity a good angular resolution is needed only in the direction perpendicular to the surface whereas in the parallel direction a large divergence can be accepted. This illumination condition is realized by a slit collimation, focusing (by e.g. elliptical mirrors) in the direction along the slit may be used to get more neutrons on a small sample area. The intensity, i.e. the reflectivity, varies over many decades from 1 below the total reflection edge dropping with the general decay of Fresnel reflectivity $\propto Q^{-4}$ into the background level in the order of 10^{-6}. Variation of Q is either performed by a variation of the inclination angle or –using a fixed angle– by variation of the wavelength. At a pulsed source the latter method with time-of-flight analysis of the wavelength is the natural choice. Even at reactor sources the time-of-fligth method is used in some cases since it leaves the geometry fixed

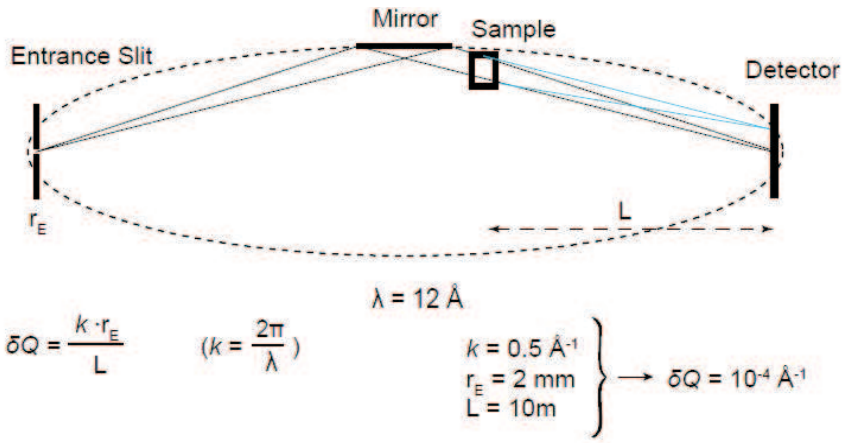


Fig. 9: *Principle of a focusing small angle neutron scattering setup which yields very high resolution in terms of the lowest observable wavevector. The focusing element is a mirror that maps a small pin-hole at the entrance of the spectrometer to a small spot on the detector. Thus very small angular deviations in the order of spot-size/sample-detector distance can be resolved. The intensity loss due to reducing the entrance pin-hole area is (partly) compensated by the large divergence that is accepted by the focusing element. The sample must be of the same area as the acceptance aperture of the focusing element, i.e. intensity is regained by illuminating a larger amount of sample.*

and there are no deformations of the reflectivity curve due to variation of the illuminated area, which occur when the sample is rotated. A recent instrument of this kind is e.g. the BioRef reflectometer at the HZB [30]. The most versatile detector at such an instrument is an area detector that allows for simultaneous measurement and immediate separation of specular and off-specular intensity. The off-specular scattering that can be analyzed in this configuration is located around the specular reflection a differing take-off angles. Since the typical reflection angles are small the structures that determine this off specular scattering are seen by the neutron beam in a projection with large contraction. Therefore in real space the corresponding scattering length density modulation are in the micrometer regime. In the perpendicular direction the off-specular scattering is sensitive to length scales that match those of a normal SANS experiment. If these modulations are also to be observed the wide divergence in the direction parallel to the surface cannot be used. Thus this special feature of a reflectometer cannot be used and therefore the experiment matches the abilities of a normal SANS instrument with *grazing incidence of the beam, GISANS*. Figure 11 shows a sketch of the scattering geometry together with example data from a microemulsion structure close to solid surface. The only difference is that now the sample is a planar surface which has to be positioned at the sample position such that the angle of incidence of the incoming neutron beam is the desired angle of specular reflection. To scan this the inclination angle of the sample must be scanned. The pattern with specular and both types of off-specular intensity is registered by the SANS area detector. The main difference to a standard SANS experiment then is the way how the data are to be interpreted, which is more complex [31].

The investigation of *magnetic layers* often needs *polarization analysis and magnetic field* at the

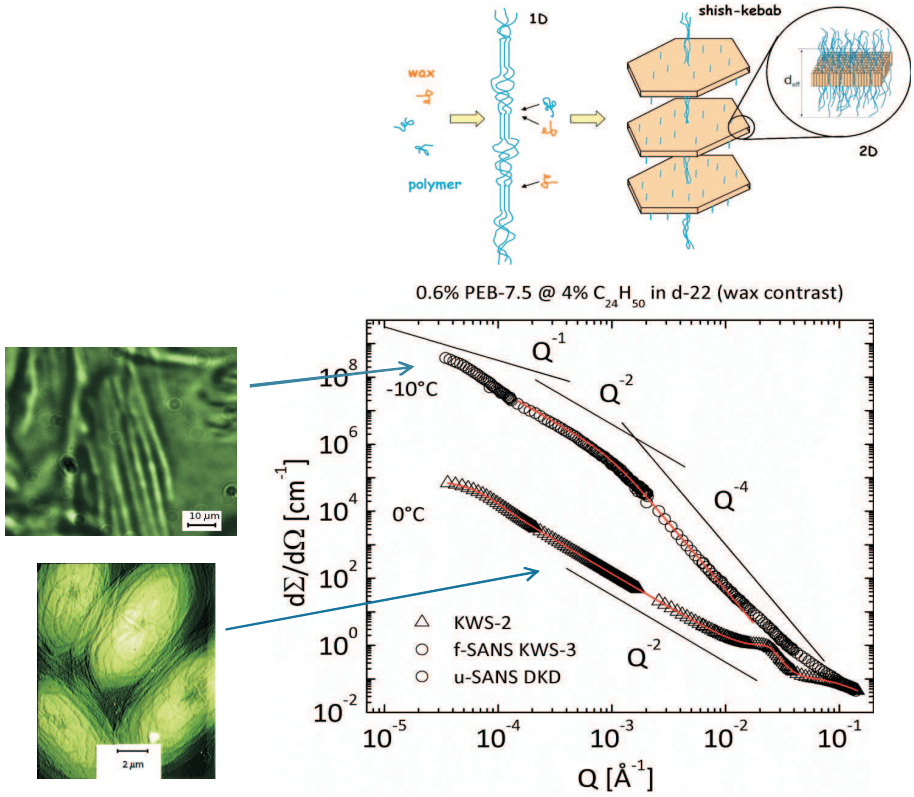


Fig. 10: *Application of a combination SANS techniques that span a very wide range in spatial resolution on a system with hierarchical structure formation. Here copolymers that are able to influence the crystallization of wax in alkanes (Diesel fuel) upon cooling are investigated in action [29]. Figure: courtesy Aurel Radulescu.*

sample. Thus there are 4 different reflectivities that can be measured by the possible combinations of incident polarization (up, down) and detected polarization (up,down). Polarization of the incident beam can be achieved using a magnetic multilayer, analysis of the reflected beam may also benefit from ^{3}He analyzer (in particular if GISANS is required). The up-down switching is performed by π-flippers which e.g. consist of a rectangular coil of Al-wire such that the neutron that enter the flipper suddenly "see" a different field orientation. The field is tuned such that a 180^0 rotation of the neutron spin is performed during its passage.
In *soft-matter and biology* an important application of neutron reflectometry pertains the structure of *liquid surfaces* covered e.g. by surfactants or similar molecules. The air liquid interface here is necessarily horizontal and cannot be tilted during scans. Thus a suitable reflectometer must supply a tilted incoming beam direction. Modern reflectometers of this type are e.g. the Platypus reflectometer at the OPAL reactor [32], the FIGARO reflectometer at ILL or at pulsed spallation sources the Liquids reflectometers at the SNS [33] and JPARC [34].

5 Spectrometers

Investigation of motions and excitation requires the analysis of energy transfer during scattering. Besides a defined and known energy before the scattering also the measurement of the neutron

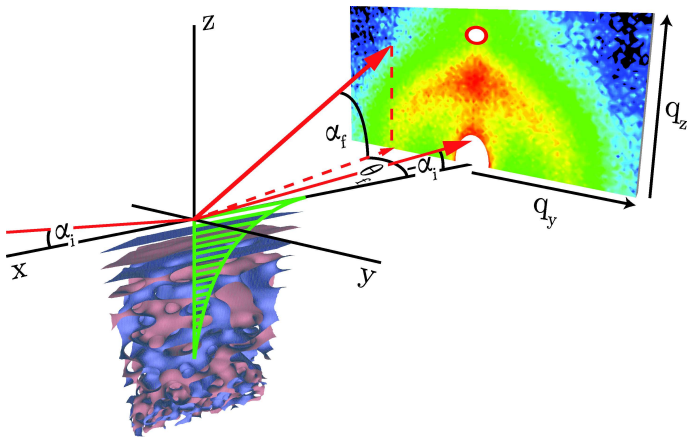


Fig. 11: *Grazing Incidence Small Angle Neutron Scattering (GISANS) geometry as it may be realized with a pin-hole SANS camera. The sketch shows –as example– a microscopic sample structure of bicontinuous microemulsion that becomes lamellar close to the surface [31]. The neutron beam enters through the material of the confining surface (e.g. Si or sapphire) at an angle below the total reflection limit. Thus the neutron enter only a small interface layer of the microemulsion in form of an evanescent wave. The penetration depth depends on the incident angle α_i and the wavelength. Figure: courtesy Henrich Frielinghaus*

energy after the scattering, i.e. in course of the detection is required. Depending on various combinations of energy ($\hbar\omega$) and momentum transfer (Q) resolution different types of neutron spectrometers are available.

5.1 Three axis spectrometer

The most obvious type of a neutron spectrometer emerges from a 2-axis diffractometer by adding a crystal Bragg reflection in the path between sample and detector. The additional rotation axis of the analyzing crystal gives this type of instrument its name: 3-axis spectrometer. The incoming wavelength is determined by the refection angle at and type of the monochromator crystal (axis 1), the incident direction is then defined using a Soller-type collimator, then the incident beam hits the sample and the scattered neutrons in a selected direction as set by the angle of the scattering arm (axis 2) and a collimator carried by it is analyzed. The analysis pertains its wavelength distribution which is determined by Bragg reflection at an analyzer crystal, the angle (axis 3) that determines the reflected wavelength. After the analyzer crystal the neutrons are counted by a detector. Eventually another collimator is used between analyzer crystal and detector. The angles may be scanned in different combinations to realize a desired path in reciprocal space $(\mathbf{Q}, \omega)$. E.g. scans at a given constant energy transfer $\hbar\omega$ may be performed along a line in $\mathbf{Q}$-space or ω may be scanned at constant $\mathbf{Q}$. In addition this may be achieved using a constant incident wavevector k_i or a constant final k_f wavenumber or some combination. If the sample is not isotropic but a crystal with excitations (e.g. phonons) that depend on $\mathbf{Q}$ a scan typically also implies rotation of the sample table respectively 3D reorientation of the sample. The instrument typically has a resolution in the order of (a few) percent. It is very versatile in the ability to set the focus to an arbitrary point in $\mathbf{Q}, \omega$ space. However, only one point at a time can be measured leading to a low data collection rate. Modern realizations of 3-axis (and similar) neutron instruments have a flexible mechanics that relies on elements as sample carrier,

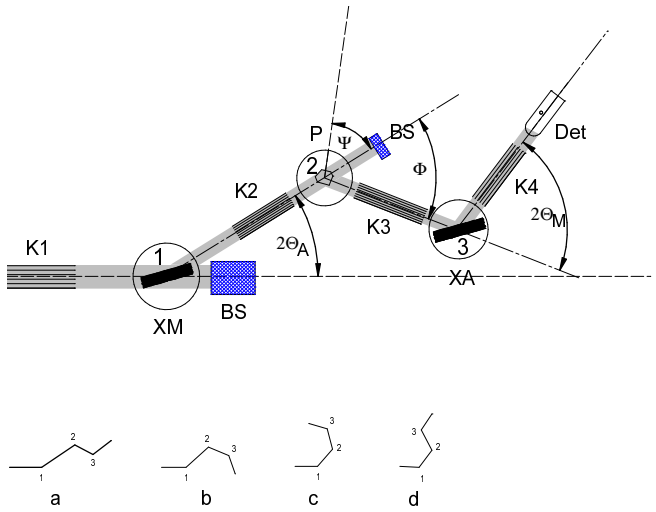


Fig. 12: *Schematics of a 3-axis spectrometer. Divergence is controlled by the collimators K1···4, the incident beam is monochromatized by a crystal XM (axis 1), the detected neutron wavelength is selected by the analyzer crystal XA (axis3). The sample is located at axis 2. Around this axis the scattering angle Φ is varied as well as the sample rotation Ψ. BS denote a beam stop. Patterns a-d show variants of spectrometer settings that can be used to sample the same $\mathbf{Q}, \hbar\omega$ point, however, the shape of the resolution is different. The latter fact is e.g. used to align an elongated tilted resolution ellipsoid to the slope of a dispersion surface.*

analyzer, analyzer detector etc. on air cushions that glide at $20\mu\mathrm{m}$ height on a marble or granite floor (Tanzboden). They are connected by levers that allow determination of the relative angles using resolvers. This enables a precise positioning of these several 100kg heavy modules. To improve the neutronic performance focusing monochromators and analyzer or analyzers that consist of crystal arrays, which reflect neutrons to an array of detectors, are employed [35]. Polarization analysis is achieved by using Heusler alloy crystals as monochromator and analyzer. For cold neutrons also magnetic supermirros and for large divergence ^{3}He cells are also in use. Three axis spectrometers are especially suited to measure excitations like phonons or magnons in single crystalline samples and to map their dispersion. An example for the measurement of magnon dispersion in copper pyrazine(pz) perchlorate measured at the PANDA 3-axis spectrometer [36] is shown in figure 13.

5.2 Time-of-flight spectrometers

Even at continuous sources time-of-flight spectrometers that utilize the stream of incoming neutrons only during about 1% of the time are well established work horse instruments. They use the flight time of neutrons over a defined path to yield information on their velocity and thus their wavelength and energy. The 3 basic different variants are the following.

- Generic time-of-flight (TOF) instrument (left part of figure 14): monochromatized beam from a crystal monochromator, chopper to create pulses (neutron bunches) in front of the sample, flight path of a few meters to detectors positioned around the sample. The final energy is inferred from the time delay between chopper opening and arrival of the scattered neutrons at the detector.

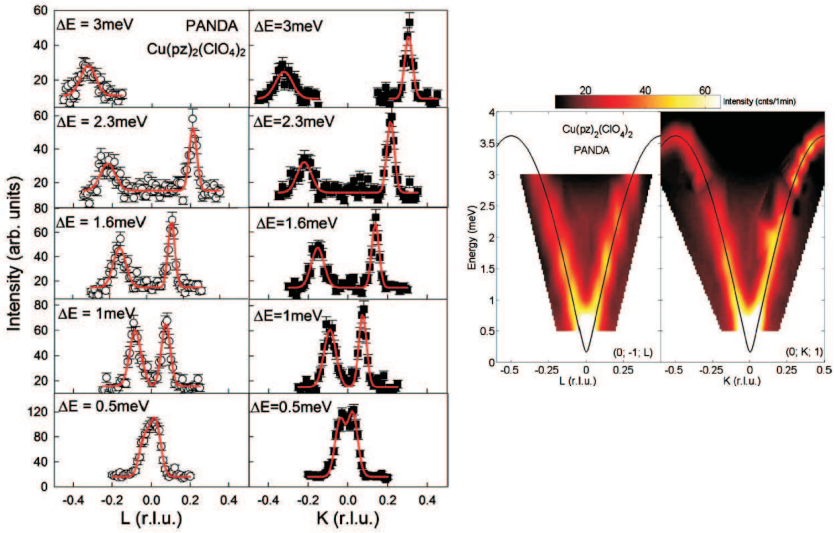


Fig. 13: *Dispersion of a gaped spin excitation at the antiferromagnetic zone center in the 2D-lattice antiferromagnet* $\mathrm{Cu(pz)_2(ClO_4)_2}$ *at low temperature (T=1.42 K) and zero external magnetic field. The left side shows two series of constant energy,* ΔE*, scans along the* $[0, k, 1]$ *and the* $[0, -1, l]$ *lines in reciprocal space. The right side compares the thus mapped excitation intensities with computed dispersion curves from linear spin-wave theory with next- and next-next- neighbor interactions. Figures from [36] with permission, copyright APS.*

- TOF-TOF instrument: instead of a crystal monochromator the incident beam path contains several choppers at a distance such that the time difference between chopper openings determines the incident neutron speed. The last chopper in front of the sample creates the neutron bunches and the detection is analog to the generic TOF.

- Inverted TOF instrument (right part of figure 14): starting from a pulsed source (chopper or spallation source) a "white" neutron beam is carried over an analyzing distance to the sample. After the sample the scattered neutrons are analyzed using an array of crystals before they reach the detectors. The final energy is fixed and determined by the analyzer crystals, the initial energy depends on arrival time of the neutrons.

The loss of intensity that is implied by the chopping of a continuous beam is largely overcompensated by the use of a large detector area that covers a very large solid angle $\Delta\Omega$ and by the fact that the whole spectrum of arrival times is collected quasi simultaneously.

State-of-the-art TOF spectrometers are equipped with position sensitive detectors that cover a huge solid angle. E.g. the IN5 spectrometer at the ILL [37] with a flight path of 4m has a detector coverage of $10\,\mathrm{m}^2$ in 100000 pixels. The energy resolution of TOF instruments is one of its defining properties, it depends on the variations of the neutron flight paths ΔL and variations in the time determination Δt. ΔL depends on the size and geometry of the sample and the variation of the exact neutron interaction location in the detectors. Both variations are in the range of several mm to cm. The time determination mainly depends on the chopper opening time, respectively the source pulse length (several 10 to 100μs). In addition the degree of monochromatization of the incoming (or in case of the inverse geometry outgoing) neutrons,

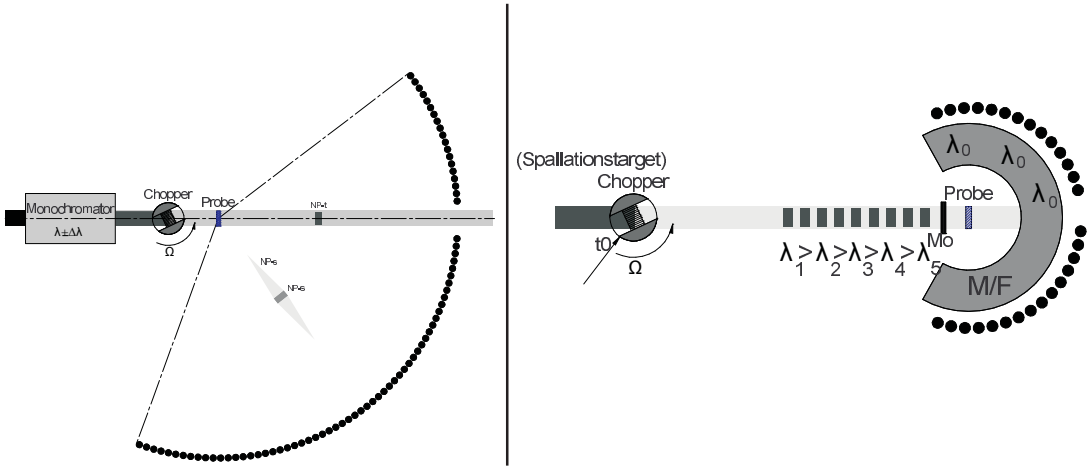


Fig. 14: *Schematics of neutron time-of-flight (TOF) spectrometers. Left: the normal geometry variant, which utilizes a monochromatizes incoming beam. A chopper in front of the sample creates a chain of short neutron bunches of uniform velocity. If the scattering is elastic neutrons from a bunch arrive at the same time at one of the detectors around the sample. Inelastically scattered neutrons vary in velocity and arrive earlier or later at the detector depending on the size and sign of the energy transfer $\hbar\omega$ during scattering. Right: inverted geometry spectrometer. Here the time-of-flight analysis of neutron energy is performed in a flight path before the sample, e.g. between sample and moderator of a pulsed source. After the sample only neutrons of a defined energy are transmitted to the detectors. The final analysis typically requires some kind of crystal Bragg reflection.*

$\Delta\lambda$ enters into the resolution width $\Delta\omega$:

$$\Delta\omega \;=\; \sqrt{(\frac{m_n L}{\hbar t^2}\Delta L)^2 + (\frac{m_n L}{\hbar t^3}\Delta t)^2 + (\frac{4\pi^2\hbar}{m_n\lambda^3}\Delta\lambda)^2} \tag{7}$$

Since $\lambda' \propto 1/v' \propto t$ close to the elastic line ($\lambda' \approx \lambda$) $\Delta\omega \propto 1/\lambda^3$. Thus the most efficient measure to increase the resolution is to increase the wavelength. This may be limited by the steep decrease of the source intensity towards long wavelength and/or by the restriction of momentum transfer range, since $Q_{\max} \propto \lambda^{-1}$. The relation for the inverse geometry TOF spectrometer are equivalent if λ is exchanged with λ'.
At larger energy transfers the nonlinear relation between time-of-flight and energy as well as the dependence of **Q** on the energy transfer becomes important (t_0, t = flight times of elastically scattered (or transmitted) neutrons and others, L flight path from sample to detector):

$$\omega(t) = \frac{m_n}{2\hbar}L^2\frac{t^2 - t_0^2}{t^2 t_0^2} \qquad\qquad Q = \frac{m_n}{\hbar}L\sqrt{\frac{t^2 + t_0^2 - 2cos(2\Theta)t_0 t}{t_0^2 t^2}} \tag{8}$$

For equal opening time of one histogramming channel the covered ω interval of one channel becomes strongly energy dependent, which has influence on the resolution at high energy transfers

and on the weighting factor connecting raw data with $S(Q,\omega)$

$$I(2\Theta,K) \propto \int_{(K-1)\Delta\tau}^{K\Delta\tau} \frac{k'}{k} 4\pi b^2 S(\mathbf{Q},\omega(t)) \frac{d\omega}{dt} dt \simeq 4\pi b^2 \frac{m_n}{\hbar} L^2 \frac{t_0}{t^4} S(\mathbf{Q}(t,2\Theta),\omega(t))\Delta\tau \qquad (9)$$

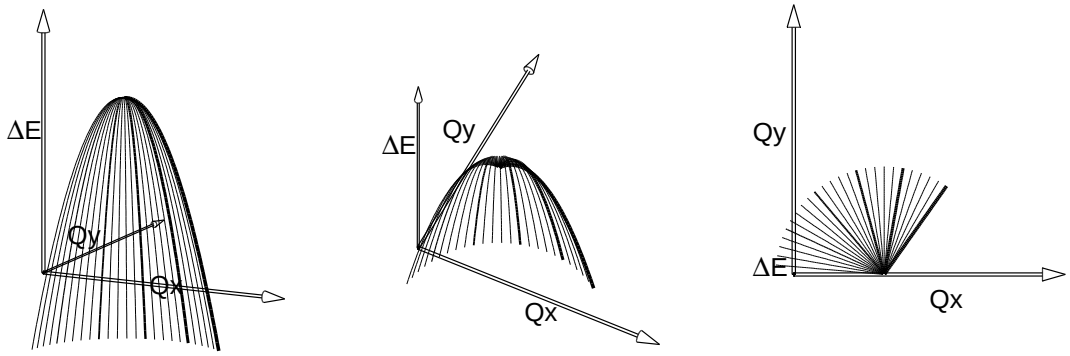


Fig. 15: *The volume in $(\mathbf{Q},\omega)$ space that is accessible in a time-of-flight neutron experiment. The indicated range corresponds to scattering angle between 5^0 and 130^0). Three different views are shown.*

The correlation between the energy and momentum transfer restricts the accessible volume in reciprocal space to the region indicated in figure 15.

The typical structure of TOF-data from a liquid is illustrated in figure 16 where the relations between arrival time an wavelength, energy transfer an Q is shown. The time-of-flight scale is proportional to the scattered, outgoing wavelength. The time corresponding to λ_0, the incoming wavelength (here 6 Å) is the position of the elastic scattering (typically the most prominent peak in the spectrum). Close to the elastic peak the momentum transfers Q of the curves for different angles are well separated and the quasi-elastic broadening of the spectra can be associated to the diffusive motion of the water molecules (tagged by the dominant incoherent scattering of the protons). Towards shorter times of flight the energy scale (dashed blue curve) and the Q scale becomes highly non-linear approaching infinity at t=0. The different spectra belong to different scattering angles. The cutoff of the inelastic spectra at short but finite time corresponds to the temperature of the sample, beyond the k_BT cutoff energy gain of the neutron becomes increasingly improbable. The energy gain part of the neutron on the other side is stretched to infinite time where the limit E_i is reached. The spectrum at that side becomes increasingly diluted in time and therefore featureless and very weak, such that a practical end has to be set. The thus established range determines the possible repetition frequency of the chopper. In contrast for a spectrometer with inverted geometry, gain and loss side are exchanged, such that excitations beyond the thermally occupied states can be excited by the neutron and detected in the spectrum. The quasi-elastic scattering around the elastic line position broadens with increasing Q and results from diffusion whereas the high energy part corresponds to vibration like motions in the liquid. The scattering intensity dominantly results from the incoherent scattering of the protons in the sample. Thus the measured spectra correspond to the self-correlation of proton motion.

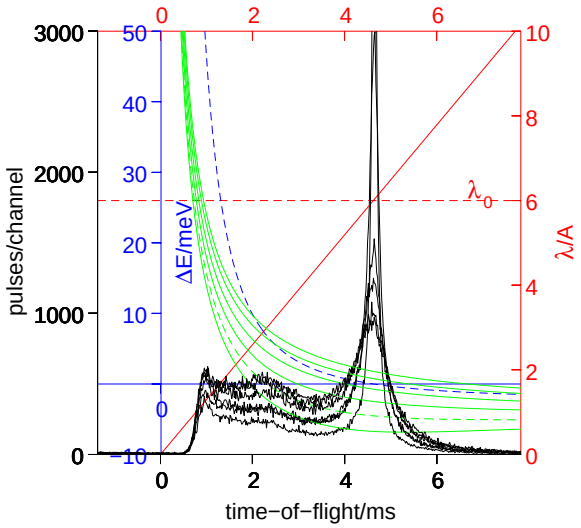


Fig. 16: *Raw time-of-flight spectra of liquid water at ambient temperature. The plot illustrates some general features of the type of information obtained.*

Another, modern example is shown in figure 17 where the position information from the detector of a disc chopper TOF spectrometer is used to directly map the magnon dispersion curve of the quasi one-dimensional antiferromagnet $(CH_3)_2CHNH_3CuCl_3$ [38].

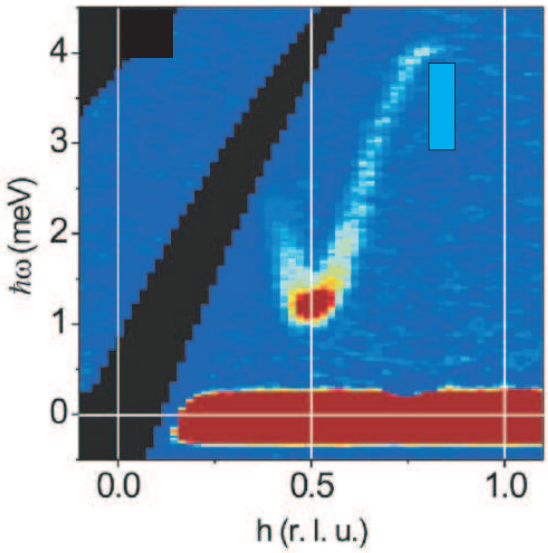


Fig. 17: *Magnon dispersion directly visible in the pixel resolved spectra from the DSC time-of-flight spectrometer at NIST, Gaithersburg [38] Figure from ref.[38], with permission, copyright APS.*

5.3 Backscattering for high resolution

To extend the energy resolution down to the μeV range energy/wavelength selection Bragg reflection from perfect crystals (e.g. the (311) reflection from Si wafers) at an angle 2Θ close to 180^0 is used. The resulting $\lambda = 2d\sin(\Theta \pm \delta\Theta)$ then only depends to second order on the divergence $\delta\Theta$. This allows to compensate the intensity loss due to the a narrow wavelength selection by allowing for a large beam divergence. At a continuous source the principle layout of a so called backscattering instrument is illustrated in figure 18. The variation of the energy transfer $\Delta E = \hbar\omega$ is achieved by the Doppler effect due to movement of the monochromator crystal in beam direction. After scattering at the sample the neutrons are analyzed by Bragg reflection backscattering from large area crystal covered "mirrors" (figure 19) that focus the intensity from a huge solid angle onto a detector. To stay in the backscattering condition theses detectors have to be close to the sample. Neutrons scattered by the sample that enter the detectors directly are discriminated from those that were analyzed by reflection at the crystal covered mirror by the longer time of flight they need for the return travel from sample position to the analyzers. At a continuous source that requires the use of a coarse chopper with about 50% opening. The chopper can be combined with mosaic crystals (PG002) that reflect neutrons from an initial white beam to the monochromator on the Doppler drive. The crystals move with the rotation of the chopper wheel and by that (if the rotation velocity is adapted) the correlation between angle and wavelength that results from the Bragg reflection condition from a mosaic crystal is modified such that the divergence of the beam is increased and at the same time the wavelength angle correlation is compensated. Then a large (focusing) monochromator crystal can be used and more intensity is available for scattering [39, 40].

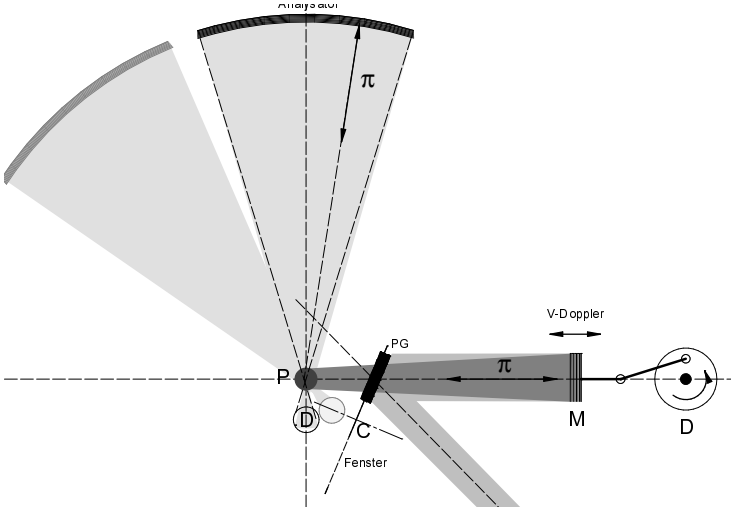


Fig. 18: *Schematics of the salient features of a backscattering spectrometer. The high resolution in the range of $\hbar\omega \simeq 1\,\mu$eV is achieved by nearly 180^0 reflection (backscattering) from perfect crystals in both the monochromator and the analyzer. The energy scan is performed by moving the monochromator in beam direction utilizing the Doppler effect to shift the neutron velocity. The intensity loss due to the narrow velocity filtering is compensated by the large solid angle covered by the analyzer segments. Each analyzer segments focuses the reflected neutrons onto one small detector.*

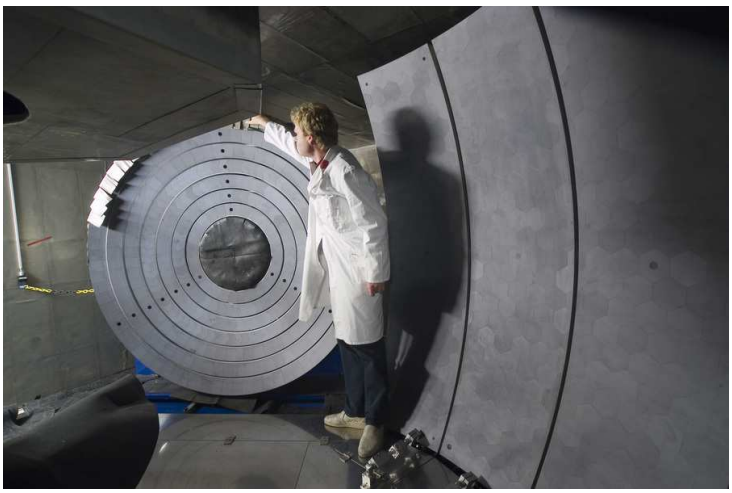


Fig. 19: *View towards the crystal covered analyzer mirrors inside the SPHERES backscattering spectrometer at the FRMII. Complete rings correspond to equal Q zones at small Q the large sectors cover the larger momentum transfers.*

As example the transition from quasi-elastic jump rotation to tunneling of CH_3 groups in theophylline with decreasing temperature is illustrated in figure 20. Due to limitations of the Doppler velocity and the finite bandwidth of the illumination via reflection from a PG002 crystal the energy width of the measured spectrum is restricted to $\pm 30 \mu\mathrm{eV}$. The energy resolution reaches values of $\Delta\hbar\omega \simeq 0.7\,\mu eV$.
This limitation is released if backscattering for analysis of scattered neutrons is used in an inverted geometry TOF instrument as e.g. the BASIS backscattering spectrometer at SNS [41]. Here the monochromatization of the beam is performed by TOF analysis with a path length of $\simeq 80\,\mathrm{m}$, the distance of the spectrometer from the moderator of the pulsed source. With these parameters BASIS achieves an energy resolution of $3 \cdots 4\,\mu\mathrm{eV}$ and a width of the spectrum of several 100 $\mu\mathrm{eV}$.

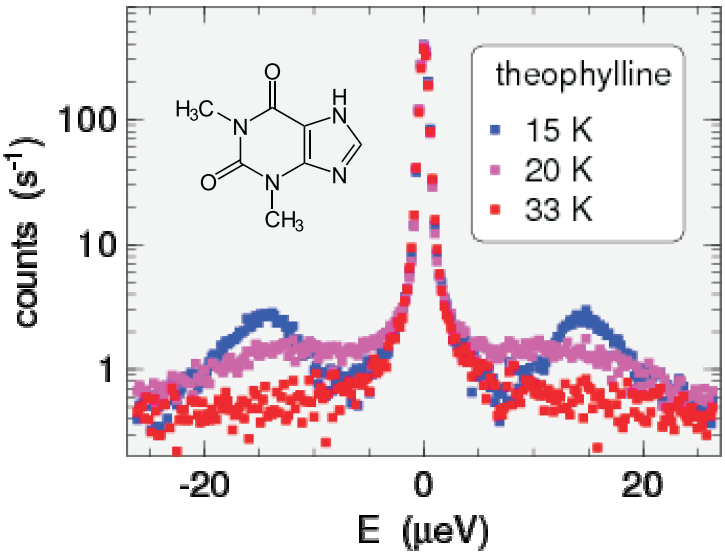


Fig. 20: *Backscattering spectrum showing tunneling lines of CH_3-groups in theophylline at low temperature and the transition to quasi-elastic rotational diffusion with increasing T.*

6 Neutron Spin-Echo Spectrometers

In general an improvement of the spectral resolution requires the narrowing of the filter transmission functions **before** (monochromator) and **after** (analyzer) the sample scattering by the desired resolution improvement factor. This implies an intensity reduction by the square of the resolution improvement factor. The attempt to increase the resolution of the TOF instruments far beyond the $\Delta E/E \approx 0.1 \cdots 1\%$ of the present realizations would lead to an unacceptable loss of intensity (as expressed in terms of detector count rate). However, it is possible to use the neutron spin directions as kind of individual stop watch pointers, thereby allowing the use of a broad velocity band (i.e. intensity) of the incoming neutron beam. The clockwork of this watch is then effected by the precession of the neutron spins in an external magnetic field [42]. [3] The restrictions affecting application are caused by the fact that the "spin-stop watch" can only be read up to an unknown integer number of complete precession turns. The scattering function $S(Q,\omega)$ is not such that there is only one defined velocity change induced but Δv is distributed according to $S(Q,\omega \approx k\Delta v)$. The detector signal is proportional to the integral of precession-angle-cosine modulated intensity contributions with weight according to $S(Q,\omega)$ corresponding to $\Delta\omega \propto \Delta v$. Therefore the signal of the NSE spectrometer is different from the TOF histograms of classical TOF-spectrometers. Instead it is proportional to the cosine-Fourier transform of $S(Q,\omega)$, i.e. the intermediate scattering function $S(Q,t)$.
A concrete setup of a NSE spectrometer is presented in the lower part of figure 21, the upper part shows the directions of the spin expectation value during the passage of a neutron through the spectrometer. Longitudinally polarized neutrons (i.e. spin expectation value parallel to the

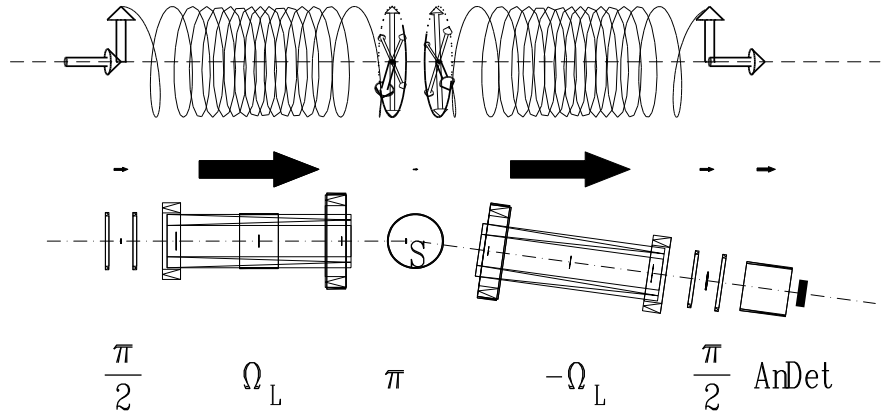


Fig. 21: *Spin rotations and setup of a generic NSE-spectrometer. Upper part: Spin rotation, middle part: magnitude of the magnetic field, lower part: schematic setup of the Jülich NSE-spectrometer.*

beam direction) enter the spectrometer from the left. In the first so-called $\pi/2$-flipper the spin is rotated such that on exit it is orthogonal to the longitudinal magnetic field of the precession

[3] It is somewhat involved to extract this analogy starting from a quantum mechanical view with spin eigenstates and eigenvalues. Implicitly we are talking about the behavior of the ensemble average of the spin vectors which obeys the "classical" Bloch equation concerning its precession in the magnetic field. As long as the kinetic energy of the neutrons is much bigger than the magnetic level splitting the classical picture is completely sufficient and much easier to visualize to understand the NSE spectrometer than a quantum mechanical treatment.

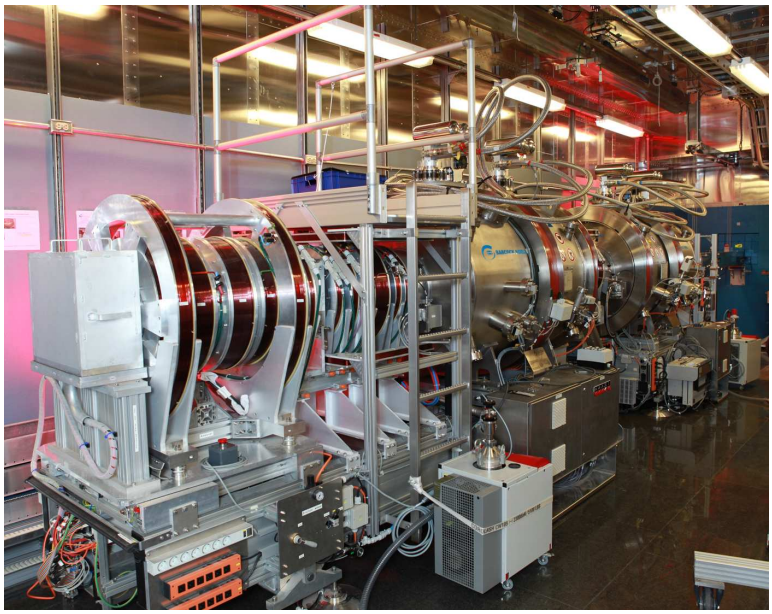


Fig. 22: *View to the NSE instrument at the SNS [43] seen from the detector end. From left (close) to the right (far) one can see the detector, the solenoids around the analyzer, the two cryostats for the superconducting main solenoids (the sample position is between these) and at the far end the blue surface of the shielding that contains the neutron guide system with frame overlap choppers. The target moderator position is 11m behind this wall.*

path. That defines the start of the "spin stop watch", immediately after the flipper a precession of the spins around the axial magnetic field begins. The precession frequency increases during the approach of the center of the main precession solenoid, where it reaches its maximum of up to a few MHz. The accumulation of precession angle continues –with decreasing frequency– until the neutrons reach the π-flipper close to the sample (S). The total precession angle at that point is:

$$\Psi_{1,2} \;=\; \frac{\gamma}{v}\underbrace{\int_{l_{1,2}} |B| dl}_{J_{1,2}} \;=\; n2\pi + \alpha \;=\; \frac{\gamma J_i}{v} \tag{10}$$

where $\gamma = 2\pi \times 2913.06598 \times 10^4 \mathrm{s}^{-1}/\mathrm{Tesla}$ is the gyromagnetic ratio of the neutrons and $|B|$ is the modulus of the magnetic induction along the primary or secondary path $l_{1,2}$.
Close to the sample (ideal: at the sample position) the so-called π-flipper is located, it rotates the spins by 180^0 around a vertical axis, thereby the total precession angle is transformed to $\Psi_1 = n2\pi + \alpha \rightarrow n2\pi - \alpha$. The parameters n and α are –according to eqn. 10 – extremely dependent on velocity and therefore very different for different neutrons in a beam with finite width of the wavelength distribution. As a consequence the spin vectors at the sample position (π-flipper) are evenly distributed on a disc orthogonal to the field direction. If no velocity change occurs during scattering at the sample (elastic scattering) each neutron enters the secondary arm of the spectrometer with unchanged velocity. The precession field and path length of the secondary arm exactly match the corresponding elements of the primary arm before the sample (π-flipper). Accordingly the precession accumulated in the secondary arm is $\Psi_2 = n2\pi + \alpha$ and the total precession angle at the second $\pi/2$-flippers is $\Psi_1 + \Psi_2 = (n+n)2\pi - \alpha + \alpha = 2n2\pi$.

I.e. all spins –irrespective of their initial velocity– reassemble at the same vertical position they had at the start point, the rotation imposed by the second π-flipper converts this back to the initial longitudinal polarization that is fully restored. The flippers limit the the two race tracks and realize "start", "time reversal" and "stop" of the "spin stop watches". The second $\pi/2$-flipper is the last element used to manipulate the spins it converts the average precession angle to a longitudinal polarization component. Since the field after the second $\pi/2$-flipper is again longitudinal, further precession does not influence the analyzed longitudinal polarization component (the stop watch is stopped!). The analyzer consists –like the polarizer– of magnetic multilayer mirrors which only reflect neutrons of one longitudinal spin state into the detector. After ensemble averaging this means that the count rate at the detector is proportional to $(1 \pm \cos(\overline{\Psi}))/2$, [4] where $\overline{\Psi}$ is the expectation value of the angle between spin and axial direction. At the point of symmetry $J_1 = J_2 = J$ it is possible to compute the velocity difference due to an energy transfer $\hbar\omega$ and to write them as series expansion for small ω:

$$-\frac{1}{\lambda^{-1}} + \frac{1}{\lambda^{-1} + \lambda(m_n/h)\omega/2\pi} \approx -\lambda^3 \frac{m_n \omega}{h 2\pi} \tag{11}$$

the detector intensity at the symmetry point ($J_1 = J_2$) is:

$$I \;=\; \eta \frac{1}{2}\left[S(Q) + \mathcal{R}(J,\lambda) \int \cos(\underbrace{\gamma J \frac{m_n^2}{h^2 2\pi}\lambda^3}_{t}\;\omega) S(Q,\omega) d\omega \right] \tag{12}$$

η contains all intensity scale factors from incoming flux to detector efficiency. The underbraced product has the unit "time", the integral in Eqn. 12 represents the cos-Fourier transform of $S(Q,\omega)$ with respect to ω, the resulting function is called **intermediate scattering function,** $S(Q,t)$ [5] From Eqn. 12 it is further recognizable that the time parameter $t = \gamma J (m_n^2)/(h^2 2\pi)\lambda^3$ depends on the third power of the wavelength λ (i.e. long wavelength $\rightarrow$ very long Fourier times). In addition $t \propto J$, i.e. mainly proportional to the current through the main precession solenoids. This current usually is the parameter used to stepwise scan the Fourier time during an experiment to get a table of $S(Q,t)$ vs. t. $\mathcal{R}(J,\lambda)$ denotes the resolution effects due to differences in precession angle for different neutron paths in the beam, which lead to a dephasing of precession angles and therefore an additional loss of echo intensity. For an ideal instrument $\mathcal{R} = 1$ would hold, in reality $\mathcal{R}(J,\lambda)$ is determined by measuring a reference sample and then used for the evaluation of experiments. The information on $S(Q,t)$ according to Eqn. 12 is contained in the ratio of the intensities at the symmetry point and the average intensity $(\eta/2)\; S(Q)$. However there are practical reasons that prevent the reliable setting of the symmetry point alone. The location of the symmetry point (i.e. phase zero current in the phase coil) is extremely sensitive to tiny variations of the magnetic environment caused e.g. by displacement of larger iron parts at neighboring instruments, movement of the crane and/or thermal displacements of coils ... . Therefore the position of the symmetry point has to be measured as well as the intensity each time. In fig. 23 the minimum of single countings is indicated, intensity must be determined for 3 points $P_1 \cdots P_3$ separated by a symmetry change corresponding to

[4] The sign in front of the cosine depends on the technical realization of polarizer and analyzer (both reflecting, transmitting, one reflecting one transmitting) and on the orientation of flippers. It may be selected by choosing the signs of the flipper currents.

[5] Strictly this is only true for a $S(Q,\omega)$ that is symmetric with respect to ω. For any practical problems however this is well fulfilled since the minute energy transfers corresponding to the NSE time scale are very small compared to $k_B T$.

a quarter precession each. From these 3 values it is possible to extract the average intensity $I(Q,0)$, the echo amplitude $I(Q,t)$ and the exact symmetry point location. This also holds if any disturbance shifted the location as indicated by the three hollow circles in the figure.

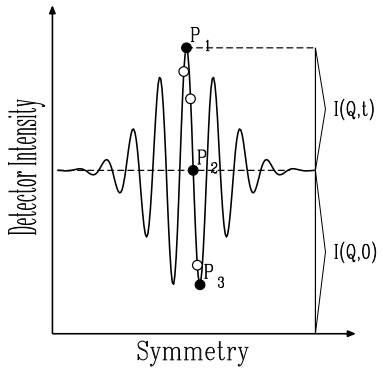


Fig. 23: *Schematic echo form, idealized.*

From the above description it follows that the NSE spectrometer measures the Fourier transform $S(Q,t)$ of the spectral part of $S(Q,\omega)$ directly. As a consequence the average count rate at the detector corresponds to half of all neutrons scattered from the sample into the solid angle of the detector (Fourier**integral**). Therefore weak spectral features are buried under the noise due to counting statistics. However the method is perfectly adapted to relaxations that are performed by most of the scattering structure. The relaxation functions are measured in the time domain directly and resolution correction consists of a division instead a deconvolution in the frequency space. One very important field for NSE investigations are "soft matter" problems. These comprise polymer melts and solutions and other complex fluids. The NSE methods opens a dynamics window in the SANS regime. Since in that regime the dynamics is determined by the balance between elastic (entropic) forces and friction and in comparison inertial forces are negligible the observed fluctuations are pure relaxations and well suited for investigation by NSE. In addition SANS scattering is coherent and usually intense compared to scattering at large angles.

The spin incoherent scattering, which frequently is the main intensity contribution in backscattering of time-of-flight spectra from organic or soft-matter samples or other hydrogen containing systems, is caused by the dependence of the scattering length on the relative orientation of nuclear and neutron spins. The fluctuating part of the scattering length due to random spin orientation contains no interference of scattering from different nuclei, i.e. the scattering intensity distributes evenly over 4π solid angle and is "diluted" accordingly. This intensity is very small compared to typical SANS intensities, in fact it is the large Q background level in SANS experiments. The dynamics of the incoherent scattering reflects the one-particle motion (self correlation). For the NSE method it is important to note that the spin-dependent scattering flips 2/3 of the neutron spins. This means that a considerable loss of polarization is encountered, only 1/3 of the neutrons contribute to the echo signal the rest is background. This 1/3 stems form the spin flipped neutrons, i.e. the echo amplitude is also inverted (negative). If coherent and incoherent scattering contributions are present this may lead to peculiar effects since the

amplitudes may cancel each other depending on their –potentially different– dynamics.

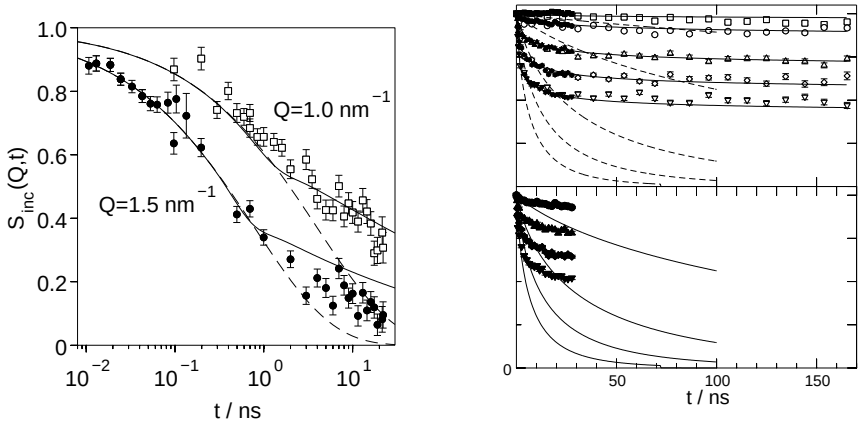


Fig. 24: *Results from neutron-spin-echo spectroscopy on polyethylene melts with height molecular weight. The left side shows the segmental diffusion at short times measured with the incoherent scattering from a protonated sample. The solid line is the expected behavior for a melt without topological constraints, the deviation at larger times indicate the first effects of a tube constraint. The right part shows plateau values where the relaxation of a chain is restricted within a virtual tube formed by the entanglements with neighbor chains. The different curves correspond to Q-values between 3 and 1.15 nm^{-1}. This experiment was done on a mixture of deuterated and protonated polymer chains. Thus the analyzed intensity corresponds to the dynamics of the single chain structure factor.*

A typical field of application for spin-echo spectroscopy is the investigation of the large scale dynamics in soft matter samples. As example figure 24 shows NSE results from a high molecular weight polyethylene melt [44]. The results reveal details of the segmental diffusion that pertain to the reptation model scenario of motion of entangled polymer chains [45]. Initially the segments diffuse without sensing the effect of other chains, at larger times the motion is hindered by topological constraints as indicated by the observation of plateaus in the $S(Q,t)/S(Q)$ decay curves [46]. These results correlate to the viscoelastic properties of polymer melts and contribute to establish a molecular explanation for their macroscopic mechanical properties.

7 Conclusion

A selection of instruments for neutron scattering has been described. The compilation cannot be complete and many further applications of thermal neutrons have been omitted, as e.g. imaging and tomography or prompt-gamma-activation analysis. Instruments at continuous sources (reactors) and those at pulsed spallation sources follow different design principles imposed by the fixed frequency pulsed nature of the spallation sources compared to the continuous ones. The common objective of the presented scattering instruments is to yield accurate data on the scattering function $S(Q,\omega)$. The variation in instrumentation originates from the variation in resolution and (Q,ω) range of interest. E.g. diffractometers integrate over ω and only structural information is obtained. Even this information is collected with different wavelength and

different angular resolution, dependent on the spatial dimensions that finally are to be probed. Inelastic instruments differ in energy resolution over a wide range of about 8 orders of magnitude spanned by the hundredths of meV range of a thermal TOF instrument down to the few neV effective resolution of a spin-echo spectrometer. A common theme for all neutron instrumentation is the optimization of intensity. This is one of the main driving forces for the variety in instrumentation. Just the necessary (and not more) resolution must be supplied in order not to reduce the available neutron flux more than necessary. Designed according to these principles a versatile and powerful suite of instruments for many purposes is found at the major neutron sources in the world. Most of these instruments are available for use by the general scientific community by application via proposal systems.

References

[1] W. Soller, Phys. Rev. **24**, 158 (1924).

[2] D. Tennant, Review of Scientific Instruments **59**, 380 (1988).

[3] J. Dash and H. Sommers, Review of Scientific Instruments **24**, 91 (1953).

[4] H. Friedrich, V. Wagner, and P. Wille, Physica B **156**, 547 (1989).

[5] E. Fermi, J. Marshall, and L. Marshall, Physical Review **72**, 193 (1947).

[6] C. Shull, Physical Review **81**, 626 (1951).

[7] C. Shull, E. Wollan, and W. Koehler, Physical Review **84**, 912 (1951).

[8] P. Boeni, Physica B: Condensed Matter **234-236**, 1038 (1997), proceedings of the First European Conference on Neutron Scattering.

[9] K. H. Andersen *et al.*, Physica B: Condensed Matter **404**, 2652 (2009), proceedings of the 7th International Workshop on Polarized Neutrons for Condensed Matter Investigations and 2nd International Symposium of Quantum Beam Science Directorate.

[10] E. Babcock *et al.*, Physica B-Condensed Matter **404**, 2655 (2009), 7th International Workshop on Polarized Neutrons for Condensed Matter Investigations/2nd International Symposium of Quantum Beam Science Directorate, Tokai, Japan, sep 01-05, 2008.

[11] G. Bruckner, A. Czermak, H. Rauch, and P. Weilhammer, Nuclear Instruments & Methods in Physics Research Section A **424**, 183 (1999), 3rd International Topical Meeting on Neutron Radiology, Luzern, Switzerland, mar 16-19, 1998.

[12] K. H. Andersen, P. M. Bentley, and L. D. Cussen, Journal of Applied Crystallography **44**, 295 (2011).

[13] T. C. Hansen *et al.*, Measurement Science and Technology **19**, 034001 (2008).

[14] H. A., H. J. P., G. O., and H. L., Zeitschrift fur Kristallographie **Proceedings 1**, 127 (2011).

[15] D. T. Bowron *et al.*, Review of Scientific Instruments **81**, (2010).

[16] H. Conrad, T. Brückel, W. Schäfer, and J. Voigt, Journal of Applied Crystallography **41**, 836 (2008).

[17] S. E. Dutton *et al.*, Journal of Physics: Condensed Matter **23**, 246005 (2011).

[18] D. Hohlwein, A. Hoser, and W. Prandl, Journal of Applied Crystallography **19**, 262 (1986).

[19] M. P. Blakeley *et al.*, Acta Crystallographica D **66**, 1198 (2010).

[20] A. J. Schultz *et al.*, Journal of Applied Crystallography **38**, 964 (2005).

[21] T. Hosoya *et al.*, Nuclear Instruments and Methods in Physics Research Section A, **600**, 217 (2009).

[22] Hammouda B., www.ncnr.nist.gov/staff/hammouda/the_SANS_toolbox.pdf .

[23] W. Schmatz, T. Springer, J. Schelten, and K. Ibel, Journal of Applied Crystallography **7**, 96 (1974).

[24] J. K. Zhao, C. Y. Gao, and D. Liu, Journal of Applied Crystallography **43**, 1068 (2010).

[25] H. Frielinghaus *et al.*, Journal of Applied Crystallography **42**, 681 (2009).

[26] E. Kentzinger *et al.*, Physica B: Condensed Matter **350**, E779 (2004), proceedings of the Third European Conference on Neutron Scattering.

[27] S. Borbely, M. Heiderich, D. Schwahn, and E. Seidl, Physica B: Condensed Matter **276-278**, 138 (2000).

[28] M. Rekveldt, Nuclear Instruments & Methods in Physics Research Section B **114**, 366 (1996).

[29] A. Radulescu *et al.*, Macromolecules **39**, 6142 (2006).

[30] M. Strobl *et al.*, Review of Scientific Instruments **82**, (2011).

[31] M. Kerscher *et al.*, Phys. Rev. E **83**, 030401 (2011).

[32] M. James *et al.*, Nuclear Instruments & Methods in Physics Research Section A **632**, 112 (2011).

[33] J. Ankner, M. Jansma, E. Blakeman, and R. Kellogg, Applied Physics A-Materials Science & Processing **74**, S1610 (2002), International Conference on Neutron Scattering, Munich, Germany, SEP 09-13, 2001.

[34] K. Mitamura *et al.*, in *Future Trends in Soft Materials Research with Advanced Light Sources*, Vol. 272 of *Journal of Physics Conference Series*, Soc Polymer Sci; Japanese Soc Synchrotron Radiat Res; Adv Softmaterial Beamline Consortium, edited by A. Takahara and K. Sakurai (IOP Publishing Ltd, Bristol, 2011), International Symposium on Future Trends in Soft Materials Research with Advanced Light Sources: Interdisciplinary of Bio- & Synthetic- Materials and Industrial Transferring, Hyogo, JAPAN, SEP 01-03, 2010.

[35] K. Lefmann *et al.*, Physica B: Condensed Matter **283**, 343 (2000).

[36] N. Tsyrulin *et al.*, Phys. Rev. B **81**, 134409 (2010).

[37] J. Ollivier, M. Plazanet, H. Schober, and J. Cook, Physica B: Condensed Matter **350**, 173 (2004), proceedings of the Third European Conference on Neutron Scattering.

[38] T. Masuda *et al.*, Phys. Rev. Lett. **96**, 047210 (2006).

[39] M. Hennig, B. Frick, and T. Seydel, Journal of Applied Crystallography **44**, 467 (2011).

[40] A. Meyer, R. Dimeo, P. Gehring, and D. Neumann, Review of Scientific Instruments **74**, 2759 (2003).

[41] E. Mamontov and K. W. Herwig, Review of Scientific Instruments **82**, (2011).

[42] *Neutron Spin Echo*, No. 128 in *Lecture Notes in Physics Vol. 128*, edited by F. Mezei (Springer, Berlin, Heidelberg, New York, 1980).

[43] M. Ohl *et al.*, Physica B-Condensed Matter **350**, 147 (2004).

[44] A. Wischnewski *et al.*, Physical Review Letters **90**, 058302 (2003).

[45] M. Doi and S. Edwards, *The Theory of Polymer Dynamics*, Vol. 73 of *International Series of Monographs on Physics* (Oxford University Press, Oxford, 1994).

[46] A. Wischnewski *et al.*, Physica B **276**, 337 (2000).

C 4 Synchrotron X-ray Beamlines [1]

Uwe Klemradt
II. Physik. Institut B
RWTH Aachen University

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

In spite of numerous attempts to diffract X-rays, seventeen years elapsed between the discovery of X-rays by W.C. Röntgen in 1895 and the first successful recording of a diffraction pattern by Friedrich and Knipping. The experiment, performed at the suggestion of Max von Laue, who also developed a detailed theoretical description, turned out to be one of the most significant scientific achievements of the 20th century. It proved not only that X-rays are capable of interference and thus possess wave characteristics, but also that crystals are periodic structures on the atomic scale, as long suspected by chemists. Therefore the famous Laue experiment can be considered the starting point of solid state physics, where (macroscopic) properties are related to the (atomic) structure. The subsequent progress in structure determination is probably best illustrated by the advent of protein crystallography (see talk E10), where the complexity has been increased to more than 10,000 atoms per unit cell, achieved nowadays routinely at large storage ring facilities usually referred to as synchrotrons.
However, synchrotron radiation can be used for far more than crystallography. Its spectrum extends from the infrared to the γ-region of electromagnetic waves, and typical applications include scattering, imaging and spectroscopy, performed on samples from practically all fields of science. As is frequently the case when novel experimental tools become available, new possibilities for precise observation are realized soon and developed into specialized techniques over time. In this overview chapter we will discuss the typical layout and similarities, but also specializations of synchrotron beamlines, concluding with two examples from the PETRA III storage ring. The focus will be on hard X-rays, although reference to soft X-rays is occasionally made. The chapter is organized using a 'downstream' approach, following the photon from the source via the monochromator to the sample and the detector.

2 X-Ray Generation

Discharge tubes were made possible by advances in vacuum technology in the 19th century, leading to the discovery of both the electron and X-rays. Although X-ray tubes have evolved over the decades, their basic design still implies a) high X-ray intensity only at characteristic wavelengths given by the target material, b) rather isotropic photon emission, c) a photon flux limited by the heat tolerance of the anode, and d) a spectrum arising from multiple scattering of electrons inside the target, which cannot be calculated analytically.
By contrast, synchrotron radiation is a very different way to generate X-rays. Being produced by highly energetic, charged particles in magnetic fields ('magnetic bremsstrahlung', see talk C2), synchrotron radiation occurs naturally in space and was on earth at first an unwanted by-product of the circular accelerators built for particle physics, which limited the available energy. After the first experimental observation in 1947, it was soon realized that the emitted light had extraordinary properties that could be used in atomic, molecular and condensed matter physics. The outstanding features are largely a direct consequence of the following two points:

1. the charged particles (typically electrons) move at highly relativistic speeds,

2. the particles are accelerated in a vacuum.

It is instructive to compare the generation of X-rays in a tube and in a synchrotron with the above list a) - d) in mind. The relativistic speeds bring in the Lorentz transformation naturally, with remarkable results for the experimentalists and their samples being in the laboratory frame.

First, the usual dipole radiation pattern in the moving electron frame is observed in Lorentz-contracted form in the laboratory frame. Together with an additional boost forward, this results in a strong, natural collimation of the emitted light cone in the laboratory frame, sometimes compared to a sweeping search light (cf. Fig. 1). The natural collimation contributes significantly to the enormous photon flux available at synchrotron beamlines, since a well-defined angle of incidence - typically required in X-ray experiments - is usually realized by slits. By comparison, simple collimator systems discard most of the photons generated by an isotropic emitter as in b).

Second, the intensity radiated is not restricted to the fundamental frequency of revolution, which is in the MHz range. The relativistic velocities together with the point-like charges give rise to harmonics up to very high order. Since the individual harmonic lines are smeared out, a continuous, broad spectrum results which extends from the infrared into the X-ray range. Therefore, a synchrotron bending magnet radiates a broad, intense spectrum, in contrast to point a) above. This allows to develop special synchrotron-based techniques which require tunable X-ray wavelengths, for example absorption spectroscopy (cf. talk F4).

Third, the acceleration of relativistic electrons by radio frequency cavities in a vacuum allows to form electron bunches of very small size (down to the micron range), resulting in an exceedingly small source size that comes very close to the ideal of a point source. This is of great importance for many experimental techniques and is described by the physical quantity 'brilliance' explained below.

Forth, it is intuitively clear that the acceleration of charges in a vacuum is a well-defined process as compared to point d) above. It leads to a clean spectrum than can be treated theoretically in full detail [1, 2]. From a practical point of view, this makes it much easier to design more elaborate equipment, like wigglers and undulators. Also the fundamental SASE (self amplified stimulated emission) process in free electron lasers (cf. talk C8), which are based on undulators, was quantitatively predicted before the first machines were built.

However, it should be mentioned that the access to synchrotron radiation is somewhat restricted, since it usually involves the submission of a proposal, with several months easily passing before an experiment can be performed. Therefore, the greatest advantage of X-ray tubes today lies in the availability of a laboratory source.

Of foremost importance is the energy E at which the synchrotron storage ring is operated,

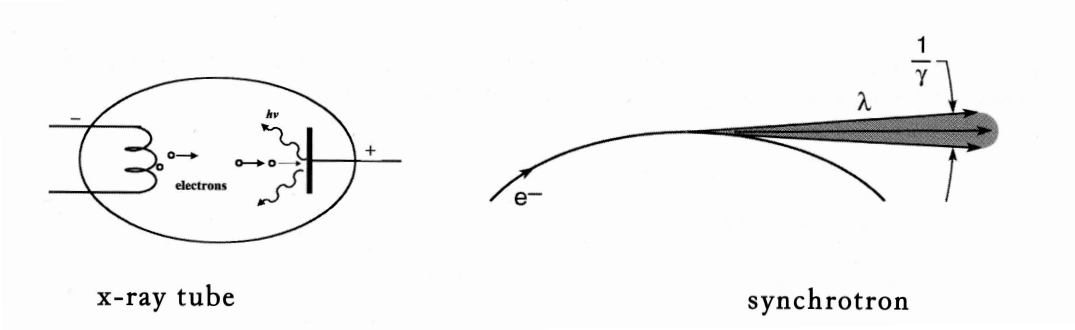


Fig. 1: *X-rays generated in a wide solid angle by a tube compared to the narrow emission angle of synchrotron radiation [3]. The angle $1/\gamma$ represents a characteristic scale for the emission in the orbit plane as well as perpendicular to it.*

typical values range from 0.7 GeV (SSLS, Singapore) to 8 GeV (SPring 8, Japan). The energy of the electrons then is almost entirely kinetic energy, since their rest mass m_0 is equivalent to only 511 keV, and the speed is only a few m/s below the speed of light according to

$$E = m\,c^2 = \gamma\,m_0\,c^2 \tag{1}$$

with

$$\gamma = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}. \tag{2}$$

Typical values of γ are several thousand for a synchrotron in the GeV range, for example $\gamma = 11742$ for a 6 GeV synchrotron. The characteristic angle $\frac{1}{\gamma}$ of the emission cone then reduces to 0.085 mrad or 0.0049º (cf. Fig. 1).
Modern storage rings are more polygons then circles, with bending magnets at the corners of the polygon and straight sections between the corners. These sections are used for so-called insertion devices, where electrons are periodically deflected from sequences of small but strong (mostly permanent) magnets. Although the emission cone is intrinsically narrow for radiation from a bending magnet, the photons are constantly emitted while the electrons follow their circular path, resulting in a large horizontal fan (e.g., in the orbit plane). By contrast, radiation from insertion devices is concentrated in a small angular range in forward direction.
The most important insertion devices are wigglers and undulators. Their basic structure is identical, depicted in Fig. 2. The periodic arrangement of magnets forces the electrons onto a sinusoidal path in the horizontal plane. An important parameter to distinguish wigglers and undulators is the deflection parameter K:

$$K = 0.934\,\frac{\lambda_U}{cm}\,\frac{B}{Tesla}, \tag{3}$$

where λ_U is the period and B the magnetic field experienced by the electrons.
In a wiggler, the N pairs of magnets create a spectrum that is equivalent to the intensity superposition of N bending magnets. However, strong magnets separated by a small gap (cm range) yield an equivalent bending magnet that is significantly stronger than the bending magnets at

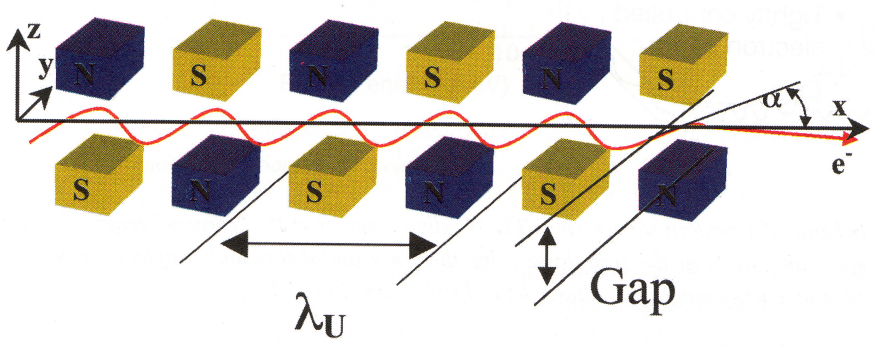


Fig. 2: *Schematic structure of wigglers and undulators. The period λ_U is typically a few cm; the strength of the magnetic field on the x-axis can be changed by varying the vertical gap between the two rows of magnets [3].*

the corners of the storage ring polygon. Therefore the wiggler spectrum is not only far more intense, but also shifted to higher photon energies. The maximum deflection angle from the x-axis is given by $\alpha = K/\gamma$, leading to a corresponding horizontal spread. Wigglers are characterized by high magnetic fields and long periods, resulting in $K \gg 1$.
Undulators are similar to wigglers except for much smaller K-values (typically $K < 1$). Their spectral and spatial radiation distributions are dominated by strong interference effects, hence undulators are not continuum sources like wigglers and bending magnets. The spectrum consists of sharp peaks at wavelengths

$$\lambda = \frac{1}{M}\frac{\lambda_U}{2\gamma^2}\left(1 + \frac{1}{2}K^2 + \gamma^2\theta^2\right), \tag{4}$$

where θ is the observation angle with respect to the undulator axis and $M = 1, 2, 3, \ldots$ indicate that also higher harmonics are generated.[2] In essence, the generation of photons with a wavelength λ on the order of 10^{-10} m in the laboratory frame is brought about by the Lorentz contraction of the undulator period λ_U (about 10^{-2} m) and the relativistic Doppler effect, yielding *two* times a γ-factor on the order of 10^4.
Since the properties of undulator radiation are based on a coherent superposition of the radiation emitted in the N periods of the device, the intensity is proportional to the square of the amplitude and hence is proportional to N^2. Undulators are frequently used at beamlines where the experiments profit most from a very intense, strongly collimated beam and need not be performed at a very specific wavelength, or when the coherence properties of the radiation are important. The tuning of the wavelength at an undulator beamline is considerably more complicated than at a wiggler beamline. Whereas at a continuum source only the monochromator needs to be changed, an undulator requires according to Eq. 4 to vary the K parameter, which is accomplished by varying the magnetic field. This is usually done by increasing the gap mechanically for shorter wavelengths and vice versa.
Finally, we want to define some useful quantities characterizing light sources. The brightness

$$B = \frac{\Delta P}{\Delta\Omega \cdot \Delta A} \quad ; \; [B] = \frac{\text{photons/s}}{\text{mrad}^2 \cdot \text{mm}^2} \tag{5}$$

is the power ΔP radiated by a source of area ΔA into a solid angle $\Delta\Omega$, with practical units as indicated. By contrast, the spectrally resolved brightness

$$B_{Spectral}(E) = \frac{\Delta P(E, \Delta E)}{\Delta\Omega \cdot \Delta A} \quad ; \; [B_{Spectral}] = \frac{\text{photons/s}}{\text{mrad}^2 \cdot \text{mm}^2 \cdot 0.1\% \,\text{bandwidth}} \tag{6}$$

is called brilliance and refers only to photons within a band pass ΔE centered at the energy E. Since relative bandwidths for crystal monochromators are on the order of $\Delta E/E = 0.1\%$ (see section 4), the brilliance is usually given in the practical units of Eq. 6.
Both quantities take into account the area of the light emitting source, which is here the cross section of the electron beam. They are fundamental figures of merit, as the product $\triangle\Omega \cdot \Delta A$ represents an invariant not affected by focusing optics (Fig. 3). It is therefore possible to produce from the same source by some suitable optics a small, divergent beam as well as a large, nearly parallel beam according to experimental needs; however the figure of merit represents a limit to what can be achieved. For further aspects, the reader is referred to talk C2.

[2] In an ideal undulator, only odd harmonics can be observed on the undulator axis.

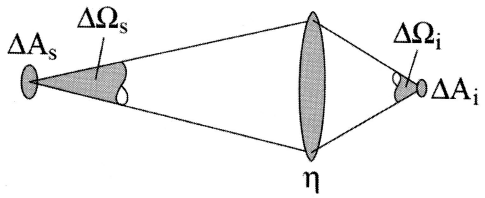


Fig. 3: *A source of finite size ΔA_S emits into a solid angle $\Delta\Omega_S$ and is imaged by an optical system with efficiency η. Under ideal conditions, the product $\Delta\Omega \cdot \triangle A$ is conserved, usually expressed via the brilliance (Eq. 6) [3].*

3 Typical Beamline Layout

The typical beamline layout found at modern synchrotron storage rings (3rd generation) is depicted in Fig. 4. All beamlines are arranged tangentially to the storage ring. From experience gathered at 2nd generation sources, the following layout has evolved:

- The optics hutch receives the incoming X-ray beam, which is modified according to the desired experimental characteristics. For example, the polychromatic beam is monochromatized here; also focusing optics are housed in this hutch.
- The experimental hutch contains the sample to be studied and a diffractometer (or other positioning devices). In case of in-situ studies, the sample environment is mounted onto the diffractometer as well, allowing to control external parameters like temperature, magnetic field etc. One or more detectors record the scattering which results from the interaction of the X-ray beam with the sample. Owing to the high level of ionizing radiation, no access during the experiments is possible, requiring remote control for each apparatus inside the experimental hutch.
- The control hutch allows the researchers to control their experiment, collect, store and transfer the data. Frequently, some basic evaluations are already performed here during the beamtime.

It should be added that hutches which contain the X-ray beam are heavily shielded in cases of storage rings producing hard X-rays. Typically, they are called lead hutches because their shielding consists of Al/Pb sandwich plates. In case of storage rings designed for soft radiation (UV and soft X-rays), lead hutches are only found at those beamlines where hard X-rays are generated by special insertion devices. The transition from soft to hard X-rays is not sharply defined, occurring at about 2 - 4 keV photon energy. At the other end of the spectrum, the soft X-ray range overlaps with the ultraviolet range, which is also called vacuum-ultraviolet (VUV) - reminding us that the radiation is so soft that it is absorbed even by air after a short distance. In these cases, the steel vessels used to transport the beam and to perform the experiments are more than sufficient to provide shielding, and hutches do not exist.
Under ideal conditions, the brilliance is conserved for a stream of photons emanating from a source and propagating down the beamline (see Fig. 3). If an experiment with a small sample requires an intense and highly parallel photon beam, a high-brilliance source like a 3rd generation synchrotron is therefore required. Also the usable photon flux for experiments which require spatial coherence is determined by the source brilliance.

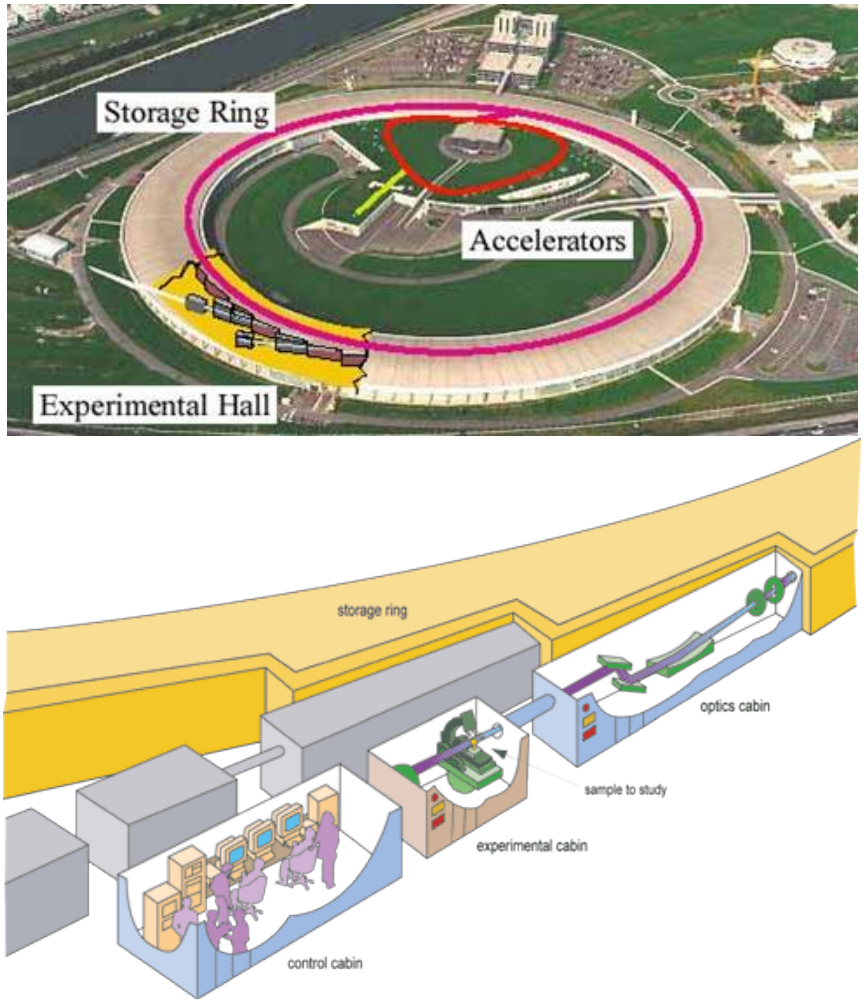


Fig. 4: *The European Synchrotron Radiation Facility, Grenoble, with a sketch of a typical beamline layout consisting of separate hutches for the optical elements, the sample interaction and the researchers controlling the experiment [4].*

4 Monochromatization

Before synchrotron radiation can be profitably applied in experiments, a monochromatization is required in many cases. There are generally two classes of X-ray monochromators:

- diffraction gratings for soft X-rays, which work well for wavelengths down to about 6 Å,
- crystal monochromators for hard X-rays.

A diffraction grating is a surface with a periodic array of geometrical features of equal shape, for example grooves ('lines'). Since a grating is a dispersive element, the combination with a slit can be used for monochromatization. As shown in elementary diffraction theory, the condition for constructive interference is

$$d\ (\sin\theta_i + \sin\theta_d) = m\lambda, \tag{7}$$

where d is the grating period, θ_i and θ_d are the angles of the incident and the diffracted beam, respectively, λ is the wavelength and m a positive integer. Monochromatization filters out a narrow band of wavelengths from a continuous distribution, with the width $\Delta\lambda$ of the band called the 'absolute resolution' of the monochromator. A high-quality monochromator provides only a small band, at the expense of the available intensity. Frequently, the parameter $\lambda/\Delta\lambda$, called 'resolving power', is used to quantify the monochromator performance:

$$\frac{\lambda}{\Delta\lambda} = m\, N_L\, (\sin\theta_i + \sin\theta_d)\,. \tag{8}$$

Since N_L is the number of illuminated lines of the grating, many lines therefore improve the resolving power. In practical setups, the grating is combined with slits and mirrors to obtain the desired performance, which is always a trade-off between high resolution and high intensity.

In the hard X-ray range, crystals are the dominant type of monochromators. The reason can be seen by the extreme case of Eq. 7 when both angles θ_i and θ_d approach 90º. In this case, the grating period d has to be on the order of the wavelength according to $d \approx m\lambda/2$, since the period cannot exceed the wavelength significantly even for high orders. For short wavelengths on the Å scale, it is technically difficult to fabricate such gratings.[3] However, the long-range atomic order of a crystal provides a natural grating with an extremely short period. The basic law of crystal diffraction is the Bragg law

$$n\lambda = 2\, d \sin\theta, \tag{9}$$

where n is the order of the diffraction, λ is the wavelength and d is the lattice spacing. Frequently, the order n is combined with the lattice spacing d to give the actual spacing probed by the X-ray wave fields at the (h k l) reflection. With d replaced by d_{hkl}, the Bragg equation now reads

$$\lambda = 2\, d_{hkl} \sin\theta, \tag{10}$$

with common factors now allowed in the values for (h k l). For the rest of this chapter, we will use this notation. A polychromatic X-ray beam impinging on a crystal will produce a characteristic set of Bragg reflections (Laue pattern). In principle, a receiving slit, together with shielding material, placed after the monochromatizing crystal can thus block unwanted wavelengths and select only the desired reflection. Suitable crystals for hard X-ray monochromators are based on a number of criteria. Of foremost importance are the 2d-values, since these should match the wavelengths of interest to give practical Bragg angles. The crystal quality is also an important parameter, since excellent single crystals have narrow reflection widths in contrast to mosaic crystals. Both types find applications in high-resolution and high-flux monochromators, respectively. Additional factors that have to be considered are thermal stability and resistance to radiation damage, especially for the first crystal encountered by the polychromatic synchrotron beam, which receives a high heat load.
The so-called 'rocking curve', a plot of the Bragg-reflected intensity as a function of the incident angle θ, is a figure of merit for each monochromator. The Bragg equation implies a δ-function, but in reality even the rocking curves for perfect crystals possess a small but finite width of a few thousands of a degree. The physics behind the finite width is that the incident X-ray wave field can only reach a finite, although large number of Bragg planes because of the finite intensity

[3]The deposition of multilayer thin films provides a possibility to create artificial one-dimensional periodic structures with a short period, which are sometimes used as low-resolution monochromators.

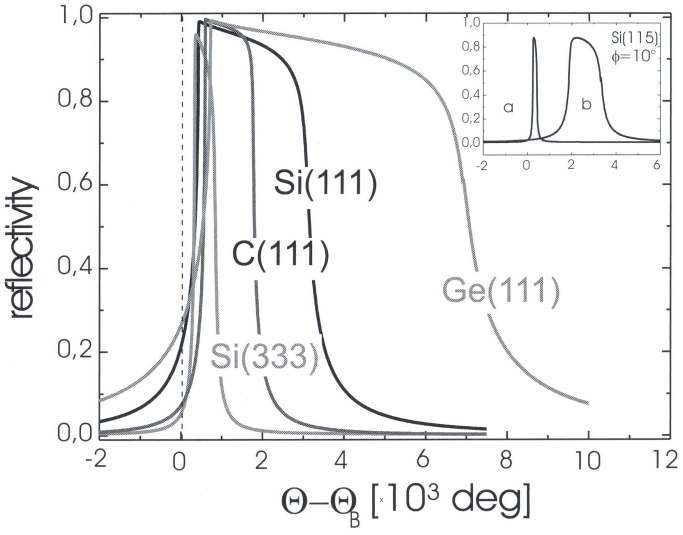


Fig. 5: *Normalized rocking curves for various [111] oriented crystals. The inset shows that the width can be reduced further when the diffracting lattice planes are not parallel to the crystal surface (asymmetric reflection) [5].*

taken out of the beam by each reflecting plane (extinction of the incident beam). The width of a Bragg reflection indexed (h k l) can be calculated for a perfect crystal in the dynamical theory of X-ray diffraction to be

$$\delta\theta_B = \frac{2\,C_{Pol}\mid \chi_{hkl} \mid}{\sqrt{\mid b \mid}\,\sin\left(2\theta_B\right)}, \tag{11}$$

where C_{Pol} is the linear polarization factor, θ_B is the Bragg angle, and χ_{hkl} the corresponding Fourier coefficient of the crystal polarizability $\chi\left(\vec{r}\right)$. The asymmetry factor b is defined as

$$b = -\frac{\sin\left(\theta_B + \phi\right)}{\sin\left(\theta_B - \phi\right)} \tag{12}$$

with ϕ being the angle between the diffracting lattice planes and the crystal surface. Typical widths of reflections are shown in Fig. 5.

There is a great variety in the design of crystal monochromators. An effective way to improve the resolving power is to use Bragg reflections from two different crystals (see Fig. 6). The highly dispersive mode is also called the (+, +) setting, since the crystals have to rotate in the same direction when changing the wavelength; consequently, the antidispersive mode is called the (+, -) setting. The latter configuration is the standard double crystal monochromator at most synchrotron beamlines, simply called DCM. It possesses the advantage that the exit beam can be kept spatially fixed when changing the wavelength by modifying the distance between the two crystals accordingly. The four-bounce monochromator depicted on the right hand side in Fig. 6 combines highest resolution with simple wavelength tuning. The multiple reflections help to suppress the tails in the single crystal reflection curves (Fig. 5), resulting in an improved resolution function.

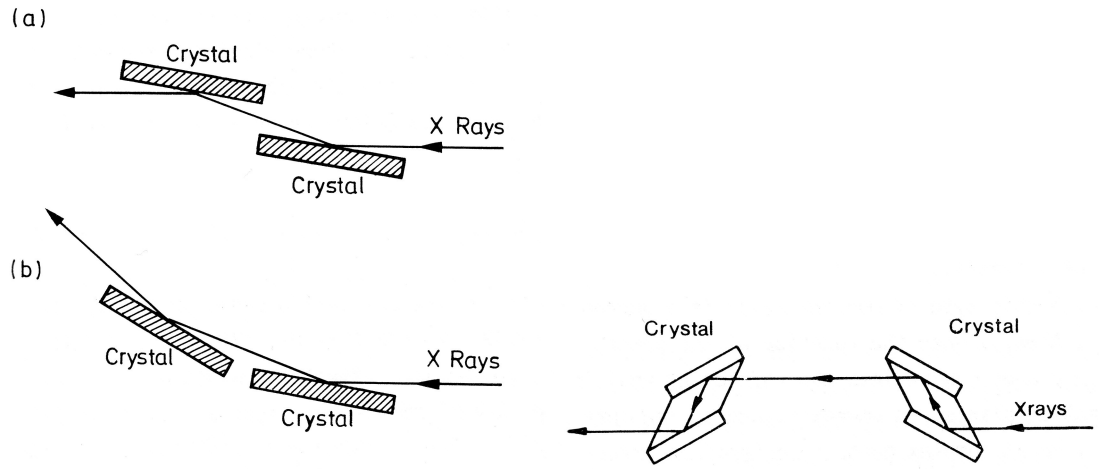


Fig. 6: *Left: Double crystal monochromators in the (a) parallel (antidispersive) and (b) highly dispersive setting. Right: Monochromator with four reflections and zero beam deviation, being a combination of the double crystal monochromators shown on the left [6].*

The resolution obtained from a single reflection is discussed in Appendix B. In essence, differentiating the Bragg equation yields:

$$\frac{\Delta E}{E} = \frac{\Delta \lambda}{\lambda} = \cot \theta_B \, \Delta \theta, \tag{13}$$

with $\Delta\theta$ being the relevant angular spread. For high resolution monochromators, this will include the divergence of the incident beam as well as the intrinsic width $\delta\theta_B$ of the reflection. Typical resolutions for good single crystals are on the order of 10^{-4} to 10^{-3}. Eq. 13 makes immediately clear that backscattering arrangements with θ_B approaching 90^o will provide the best resolution.

DuMond diagrams are helpful to visualize the band selected from a continuous spectrum due to the finite angular acceptance of a particular Bragg diffraction (Fig. 5). Whereas the Bragg equation couples λ and θ to each other, the idea here is to view them as independent because of the finite reflection width given by Eq. 11, and hence display the wavelenghts transmitted by a monochromator in the λ-θ -plane. For simplicity, the transmission function is approximated by a rectangular function, so that the basic plot in a DuMond diagram consists of a graph of the $\sin$-function from Eq. 10, but modified to have a finite width (see Fig. 15) [7].

The different resolutions resulting from the (+, +) and (+, -) settings for two crystals can easily be understood using DuMond diagrams. In the dispersive (+, +) setting, the transmission curve of the second crystal must be plotted in the opposite direction. The small overlap area indicates that only a small range of wavelengths leaves the monochromator. The excellent resolution of this arrangement can be used for laboratory tubes to even separate the $K\alpha_1$ and $K\alpha_2$ lines, as indicated in Fig. 7. By contrast, the parallel (+, -) setting implies two parallel graphs with a large overlap, i.e. all wavelengths diffracted by the first crystal are diffracted by the second crystal as well, which results in far greater flux after the monochromator.[4]

[4]Of course, in a strict sense the (+, -) setting is non-dispersive only if both crystals are equal and the same Bragg reflection is used.

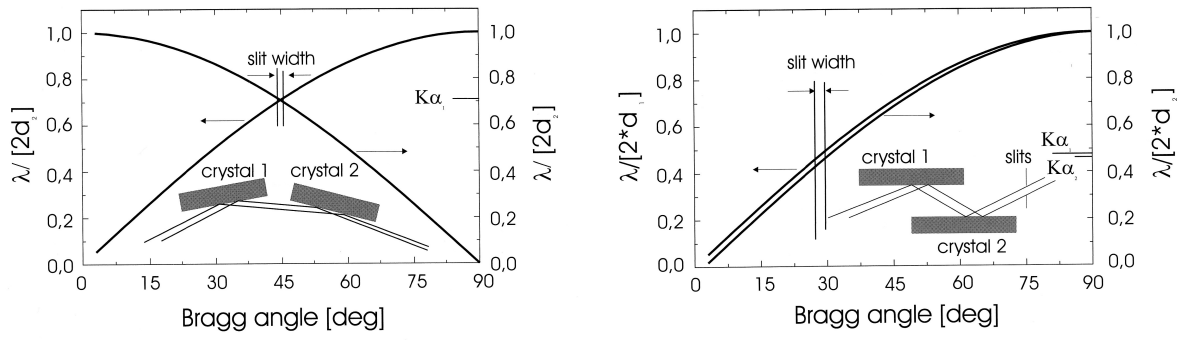


Fig. 7: *DuMond diagrams and corresponding crystal arrangements for double crystal monochromators in the antidispersive and dispersive setting [5].*

It should be noted that the broad spectrum of polychromatic synchrotron radiation leads generally to the unwanted admixture of higher-order wavelengths like $\lambda/2$, $\lambda/3$ etc. in the monochromatized beam if no special measures are taken. When Si or Ge (111) reflections are used for monochromatization, the important second-order harmonics are already suppressed because of the structure factor of the diamond lattice (e.g., (222) is a forbidden reflection with zero intensity); unfortunately, this does not provide a general method for higher-order rejection. Owing to refraction effects, the higher-order harmonics possess absolute Bragg angles slightly but systematically shifted away from the fundamental reflection. This can be used to suppress higher harmonics by slightly rotating the second crystal of the DCM (on the order of a thousand of a degree), thereby moving the center of the band pass away from the harmonics. Another popular means of suppressing harmonics is the reflection from mirrors using an appropriate angle of incidence. The wavelength-dependent total reflection of X-rays under grazing incidence up to a critical angle (see talk D2) makes mirrors effective low passes; in addition they can also serve to achieve focusing in one or two planes with an appropriate curved shape.

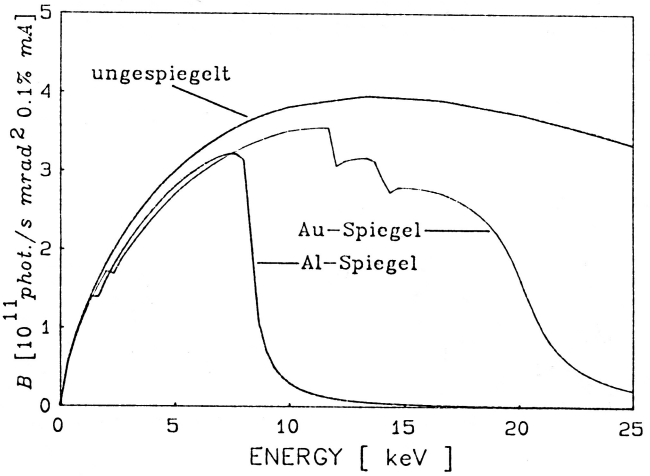


Fig. 8: *The total reflection of mirrors under grazing incidence provides an effective low pass frequently used for the suppression of higher harmonics. The topmost curve is an unfiltered bending magnet spectrum. The structures in the curve of the Au mirror result from absorption edges [8].*

5 Sample Interaction

Standard instruments for angle-dispersive measurements at fixed wavelength are so-called four- or six-circle diffractometers with sufficient degrees of freedom to scan or map those parts of reciprocal space which are of particular interest in the problem to be investigated. In the simplest case, encountered in many laboratory setups, only two circles are present: the sample rotation (usually called θ or ω), and the detector rotation (usually called 2θ, although this degree of freedom is independent from the sample rotation and is only twice the detector angle in certain scan types). Both rotations are restriced to the plane of incidence.
Four- and six-circle diffractometers possess additional degrees of freedom for sample positioning, which open up more possibilities in accessing reciprocal space. For example, the six-circle diffractometer combines the features of a four-circle and a so-called z-axis diffractometer [9]. It is useful in surface sensitive scattering, where a grazing angle of incidence (the control of which is important for depth resolution) must be kept constant during a scan and independent from the Bragg angle of the reflection to be measured.
The degrees of freedom for a six-circle diffractometer are (cf. Fig. 9):

- δ – detector arm rotation
- ω – rotation of the sample circles
- χ – sample tilt
- φ – sample rotation
- α – rotation of the entire diffractometer
- γ – out-of-plane detector rotation

There are also special setups for energy-dispersive measurements performed at fixed angles, usually either because there is no space for large angular variations (for example in high-pressure diffraction with anvil cells) or because of time gains (parallel measurements of a large range of q-values). The interested reader is referred to the literature for details [10].

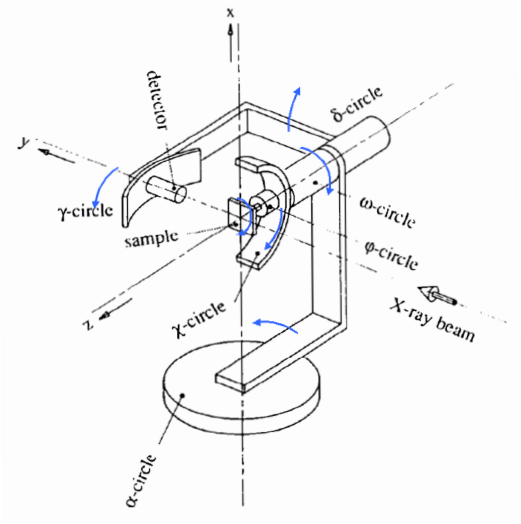


Fig. 9: *Degrees of freedom for a classical six-circle diffractometer [9].*

6 Analysis of the Scattered Beam

The scattered beam can be analyzed in a variety of ways with respect to the required angular or energetic resolution. In the simplest case of a scintillation or gas counter used as a point detector, the angular divergence of the detected radiation is typically reduced by a slit or a collimator system (for example Soller slits), thus providing angular resolution in scans. If reciprocal space maps are measured, a good resolution is required in two directions [5]. High-resolution measurements frequently involve so-called analyzer crystals. Quite analogous to monochromator crystals, one or more analyzer crystals can be positioned behind the sample, deflecting the scattered beam into the detector. They can be viewed as an extreme form of slits with a width given by Eq. 11. However, they are more than that, since the scattered radiation has to fulfil the Bragg equation, resulting in a very effective suppression of inelastic background (for example, fluorescence or Compton scattering). Analyzer crystals are sometimes bent to achieve higher intensity if a relaxed angular resolution is acceptable.

The advent of very bright undulator radiation has also brought possibilities for the measurement of extremely weak *inelastic* scattering. This provides experimental access to elementary excitations in solids, but also to Mössbauer studies with synchrotron radiation (see talk D9). A schematic setup for inelastic measurements is shown in Fig. 10. The magnitude of the problem becomes clearer if one realizes that, for example, in such experiments phonon energies in the 1 - 100 meV range are to be probed with 20 keV photons. Extreme resolutions in the meV or sub-meV regime can, according to Eq. 13, only be reached with analyzer crystals in backscattering geometry, using highly indexed reflections like Si (9 9 9) etc. At such high resolutions, energy scans are not performed any more by rotations, but by varying the relative temperature of the monochromator and analyzer crystals with mK precision, since the lattice constant enters the reflection width according to Eq. 21. For more details on inelastic X-ray scattering and the various spectrometer designs, the reader is referred to the literature [11][12] .

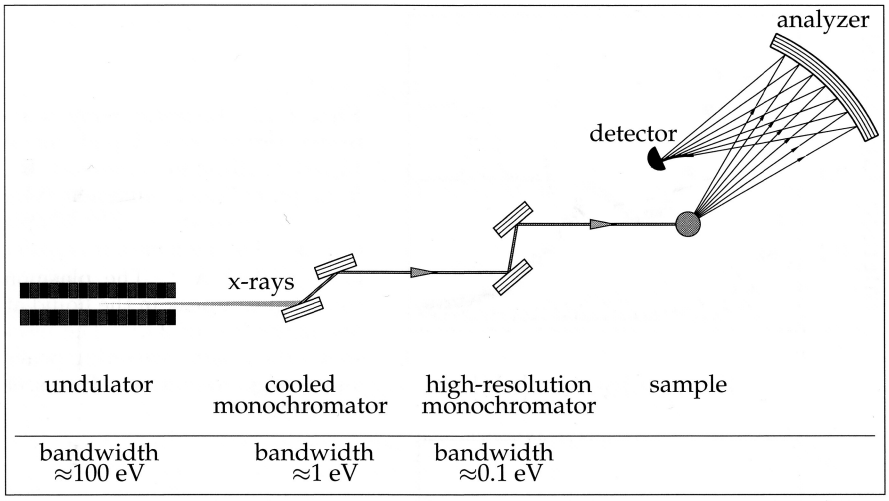


Fig. 10: *Measurements of inelastic processes require a cascade of monochromator and analyzer crystals. The highest resolution is achieved in backscattering geometry with a flat analyzer crystal, whereas bent crystals provide more intensity [11].*

7 Specialized Beamlines

A modern synchrotron facility hosts many beamlines, all of them tailored to specific tasks and receiving radiation from insertion devices which exactly match their purposes. The following list, being certainly not exhaustive, is intended to give merely an impression of the degree of specialization:

- crystallography (powder diffraction, protein crystallography, high pressure)
- topography
- small angle scattering (SAXS, GISAXS)
- surface diffraction and grazing incidence
- absorption spectroscopy (EXAFS, XANES, dichroism)
- imaging and microscopy
- inelastic scattering (Compton, high resolution, nuclear resonance)
- coherence applications
- photoelectron spectroscopy.

To provide an illustration, the beamline layout of the PETRA III facility in Hamburg is depicted in Fig. 11. PETRA was a storage ring originally designed for particle physics experiments. When converted into a 3rd generation synchrotron radiation source in 2007, one octant of the ring was equipped with 14 beamlines with more than 40 instruments.

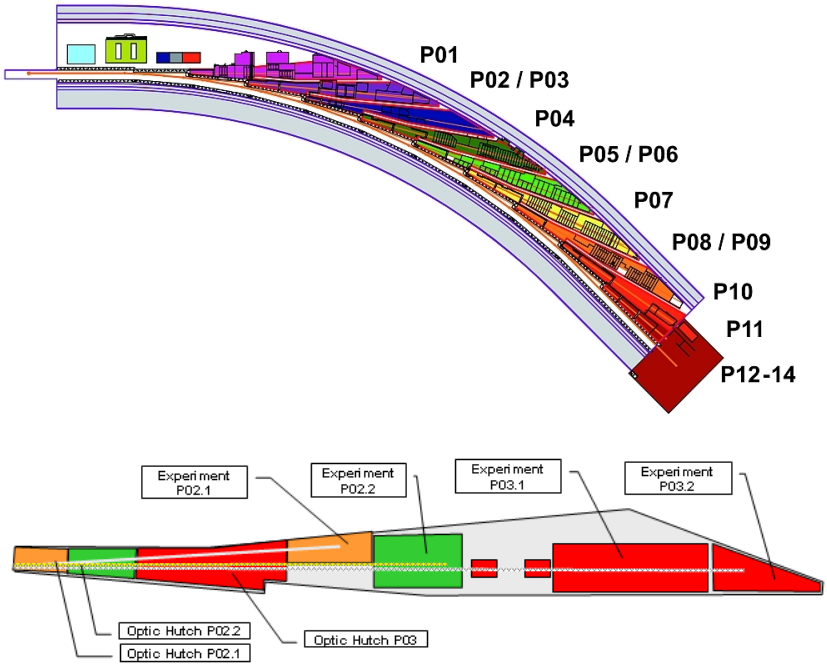


Fig. 11: *Top: Sketch of the PETRA III beamline layout. Bottom: Enlargement of sector 1 and 2 hosting the beamlines P02 and P03, each with two experimental endstations. The X-ray beam enters from the left [13].*

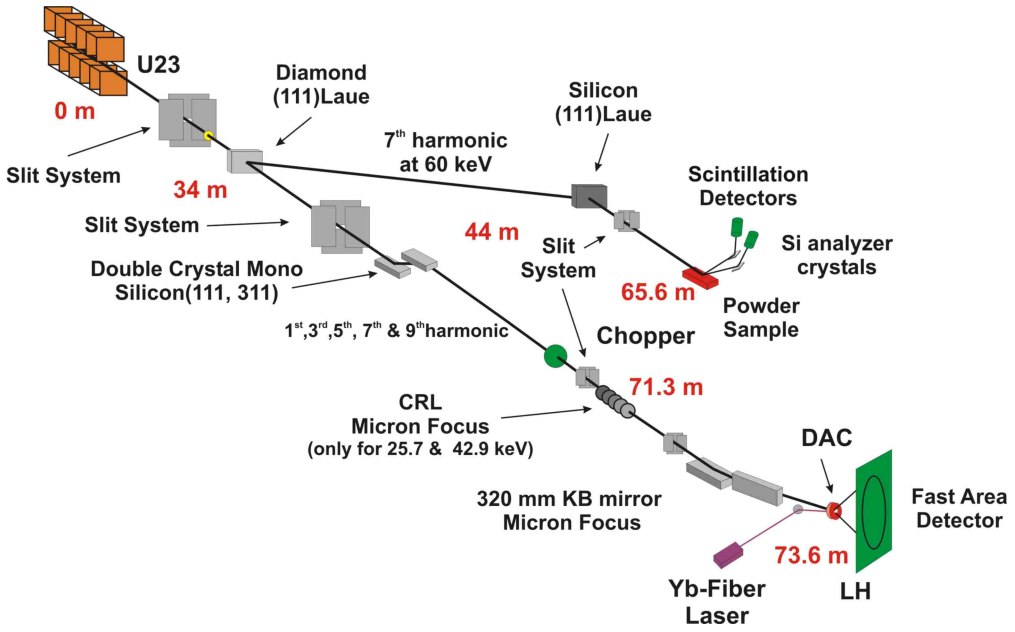


Fig. 12: *Layout of the PETRA III beamline P02 [13].*

The beamlines in sector 1 and 2 are designed for inelastic X-ray scattering and nuclear resonant scattering techniques (P01), very hard X-rays (P02) and small-angle X-ray scattering with micro- and nanofocus capabilities (P03). We will have a closer look at P02 and P03 in the next paragraphs. Other beamlines deal with soft X-rays, imaging, coherence and life science applications, including small-angle scattering and macromolecular crystallography.

A schematic of beamline P02 is shown in Fig. 12. The beamline is designed to provide relatively high-energy X-rays with high brilliance and small focus for diffraction experiments. One endstation operates at 60 keV, using the 7^{th} harmonic of the undulator. Its purpose is high-resolution powder diffraction at ambient pressure and variable temperature. The second endstation is designed for micro powder and single crystal diffraction at extreme conditions, for example simultaneous high pressure and variable temperatures. Both endstations are capable of time-resolved diffraction, useful to investigate non-equilibrium processes [13].

The high brilliance of PETRA III also allows to realize extremely focused beams, leading to high resolution both in real and reciprocal space. One example is the micro- and nanofocus small- and wide-angle X-ray scattering beamline MINAXS, designed to provide corresponding beams with very high intensity. The beamline consists of separate micro- and nanofocus endstations in the hutches P03.1 and P03.2, respectively, designed in collaboration with external groups. The microfocused beam is formed by compound refractive lenses (CRL) and has a variable size, typically 10 x 10 microns^2. The size of the nanofocused beam is about 100 x 100 nm^2, obtained by X-ray waveguides. The beamline is designed to operate between 8 and 23 keV; to suppress higher harmonics, a planar double-mirror is used. Since the two endstations use the same monochromatized synchrotron beam, they can only be used alternately [13].

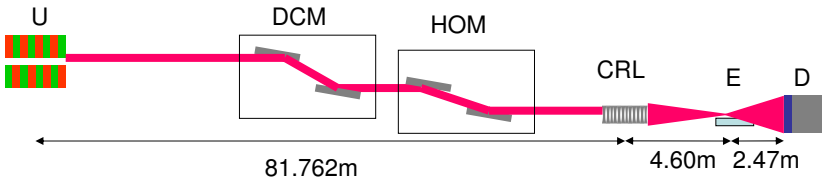


Fig. 13: *Sketch of the microfocus endstation of beamline P03. HOM stands for higher order suppression mirrors, CRL for compound refractive lenses. The sample is positioned at E, the detector at D [14].*

We want to conclude with an example for the research possibilities with microbeams. In a recent study performed at the microfocus endstation (see Fig. 13), the nanostructuring and cluster formation of Au nanoparticles on top of a thin polymer film were investigated by Roth et al. [14]. The experimental method used was grazing-incidence small angle X-ray scattering (GISAXS), depicted in Fig. 14, which is well suited for the investigation of structural inhomogeneities at and near surfaces. For this type of experiment, 2D detectors are used which register in-plane and out-of-plane scattering simultaneously. The real-time experiment followed the temporal evolution over two orders of magnitude to about 1000 s. The huge amount of data gathered with 2D detectors during such experiments requires new approaches for data evaluation [15]. As a result, four different stages of nanocomposite formation were identified, ranging from diffusion and roughness increase to layer build-up and compaction.

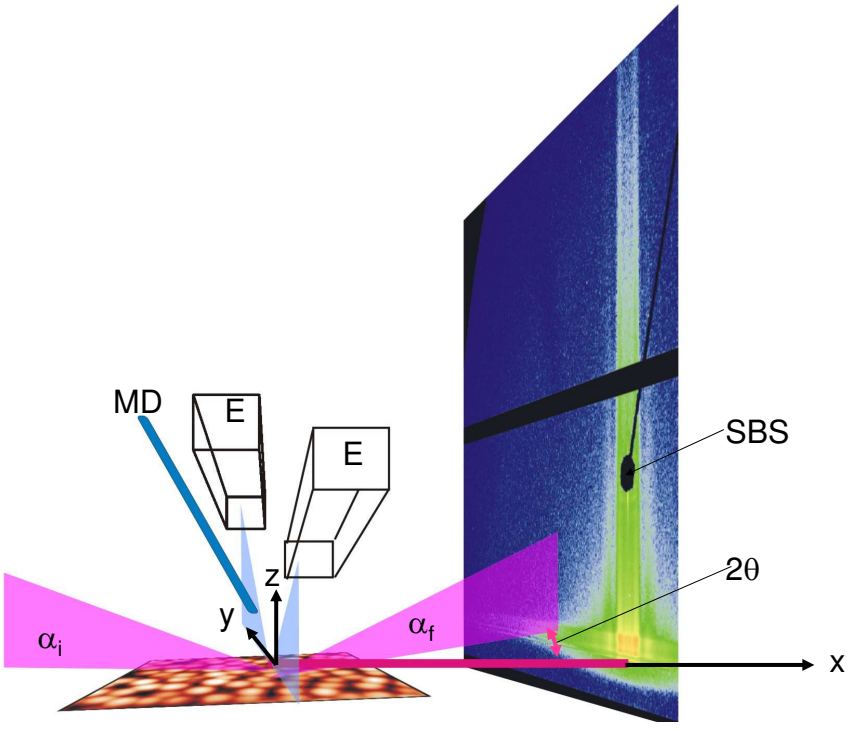


Fig. 14: *Geometry of a small angle scattering experiment under grazing angles α_i and α_f. The sample is additionally characterized by imaging ellipsometry (E); a retractable micropipetting system allows droplet deposition (MD). The specular beamstop is denoted as SBS [14].*

Appendices

A Energy-Wavelength Relation for Photons

The relation between energy and wavelength is for photons

$$E = \frac{hc}{\lambda} \tag{14}$$

or in practical units

$$E(keV) = \frac{12.398}{\lambda(\mathring{A})}. \tag{15}$$

Differentiating Eq. 14, one obtains

$$dE = -\frac{h\,c}{\lambda^2}\,d\lambda, \tag{16}$$

where the minus sign signifies that an increase dE in the energy is related to a decrease $d\lambda$ in the wavelength. For a finite but small energy interval ΔE , one obtains therefore for the relative resolution

$$\frac{\Delta E}{E} = \frac{\Delta\lambda}{\lambda}, \tag{17}$$

where the minus sign has been dropped since $\Delta\lambda$ is usually also defined to be a positive quantity.

B Flux and Resolution in a DuMond Diagram

We start with the Bragg equation in the form of Eq. 10

$$\lambda = 2\,d_{hkl}\sin\theta. \tag{18}$$

In a continuum of polychromatic X-rays, any small angular spread $d\theta$ is related to a corresponding spread in wavelength $d\lambda$ obtained by taking the differential of Eq. 18:

$$d\lambda = 2\,d_{hkl}\cos\theta\;d\theta, \tag{19}$$

resulting in the spectral resolution

$$\frac{d\lambda}{\lambda} = \cot\theta\;d\theta \tag{20}$$

by dividing both expressions. These are our basic equations, which are now applied to two different situations.

First, we look at the intrinsic wavelength spread $\delta\lambda$ associated with the intrinsic width $\delta\theta_B$ of a Bragg reflection indexed (h k l). Inserting Eq. 11 in Eq. 19 and using the identity $\sin(2x) = 2\,\sin x\,\cos x$ yields

$$\delta\lambda = 2\,d_{hkl}\cos\theta_B\;\delta\theta_B = d_{hkl}\,\frac{2\,C_{Pol}|\chi_{hkl}|}{\sqrt{b}\,\sin\theta_B}. \tag{21}$$

Second, we look at the flux resulting from polychromatic radiation with an angular divergence $\Delta\theta$. The flux is proportional to the shaded area in Fig. 15.

$$F \sim \delta\lambda \cdot \triangle\theta \tag{22}$$

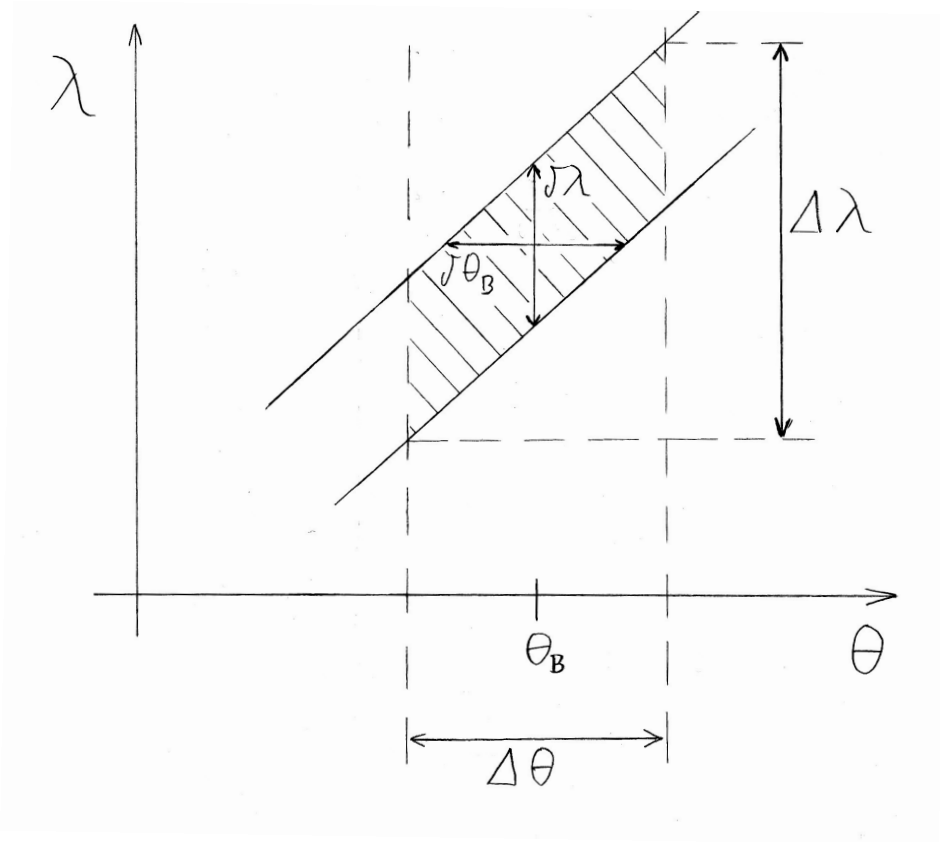


Fig. 15: *DuMond diagram for a single crystal monochromator, showing only a small part of the* sin*-curve. The photon flux through the monochromator is proportional to the shaded area.*

We now want to calculate the spectral resolution for a situation where both the angular divergence $\Delta\theta$ and the intrinsinc spread $\delta\theta_B$ contribute significantly to the total wavelength spread $\Delta\lambda$ using Eq. 20:

$$\frac{\Delta\lambda}{\lambda} = \cot\theta_B \ (\Delta\theta + \delta\theta_B)\,. \tag{23}$$

With the help of Eq. 21, we obtain the final result

$$\frac{\Delta\lambda}{\lambda} = \cot\theta_B\,\Delta\theta + \frac{C_{Pol}|\chi_{hkl}|}{\sqrt{b}\,\sin^2\theta_B}. \tag{24}$$

References

[1] D.D. Ivanenko, I. Pomeranchuk, Phys. Rev. **65**, 343 (1944).

[2] J. Schwinger, Phys. Rev. **70**, 798 (1946).

[3] 38thIFF Spring School, chapter B.2 (Forschungszentrum Jülich, 2007).

[4] www.esrf.eu, accessed Dec. 2011.

[5] U. Pietsch, V. Holý, T. Baumbach, *High Resolution X-Ray Scattering: From Thin Films to Lateral Nanostructures*, chapter 1 (Springer, Berlin, 2004).

[6] C. Kunz (ed.), *Synchrotron Radiation: Techniques and Applications*, chapter 3 (Springer, Berlin, 1979).

[7] J.W.M. DuMond, Phys. Rev. **52**, 872 (1937).

[8] 23rdIFF Spring School, chapter 20 (Forschungszentrum Jülich, 1992).

[9] M. Lohmeier, E. Vlieg, J. Appl. Cryst. **26**, 706 (1993).

[10] D.M. Mills (ed.), *Third-Generation Hard X-Ray Synchrotron Radiation Sources: Source Properties, Optics, and Experimental Techniques*, (Wiley, New York, 2002).

[11] Y. Shvyd'ko *X-ray Optics: High-Energy-Resolution Applications*, chapter 1 (Springer, Berlin, 2004).

[12] H. Sinn, J. Phys.: Condens. Matter **13**, 7525 (2001).

[13] hasylab.desy.de/facilities/petra_iii, accessed Jan. 2012.

[14] S.V. Roth, G. Herzog, V. Körstgens, et al., J. Phys.: Condens. Matter **23**, 254208 (2011).

[15] Directly Programmable Data Analysis Kit DPDAK, ww.desy.de/~beneckeg/dpdak

C 5 X-Ray & Slow Neutron Detectors

G. Kemmerling
Zentralinstitut für Elektronik
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

This lecture is about radiation detectors, which are used for the detection of X-rays and slow neutrons. After a short review of the relevant interactions of X-rays and neutrons with matter, the basic operating principles of these detectors and their specifics for X-ray and slow neutron detection will be described. However, neither all details of these detectors nor their application in different research area can be covered within this lecture. The books, listed in the references serve for this purpose. While references [1]-[3] give a comprehensive introduction and overview of the different types of radiation detectors, references [4] and [5] are specialized to semiconductor and scintillation detectors.

2 X-Ray and Slow Neutron Interactions with Matter

The common detection principle of radiation detectors is based on the ionization of the detector material, which gives rise to an electrical signal to be measured by signal processing electronics. As the ionization is due to the interaction of the particle radiation with matter, these interactions need to be known for an understanding of the radiation detectors. The following section gives a brief review of the interactions of X-ray and neutron radiation with matter, which are relevant for radiation detectors.

2.1 The interactions of X-Ray and Gamma-Radiation

For radiation detectors, there are three important interactions of X-rays or gammas with matter (see Fig. 1).

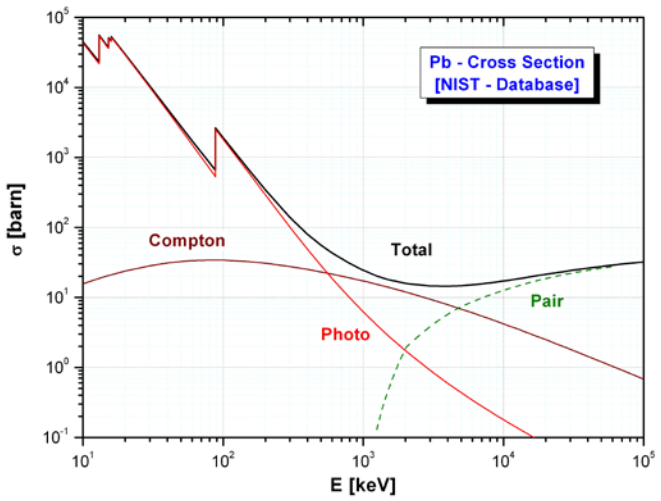


Fig. 1: Cross sections of the important interactions of X-ray and gamma radiation with matter, giving lead as example.

At X-ray energies up to about 100 keV, the photoelectric effect is dominant. There, an incident photon is absorbed by an atom and an electron is ejected. Due to energy and momentum conservation, the atom takes a part of the photon momentum. Thus, always the possible strongest bound electrons absorb the photon. As a result of the de-excitation of the

atom, a characteristic fluorescense photon or, especially at higher Z, Auger electrons are emitted. The higher the X-ray energy becomes, the electron appears to be less bounded for the photon. Therefore, there is a strong decrease of the cross section with increasing energies and a competitive effect, the Compton scattering, becomes important up to energies of 5 MeV. Compton scattering is incoherent scattering on a quasi-free electron of an atom accompanied by an energy and momentum transfer from the photon to the electron. At energies above 1.02 MeV, a third process called pair production gets relevant. Pair production is the process of creation of an e^+/e^--pair, mostly in the field of the nucleus. As the e+ is not stable in matter, it annihilates after stopping into two 0.51 MeV photons.

2.2 Converter Materials for Slow Neutrons

Slow neutrons are understood to have energies of up to some 100 meV. As neutrons don't have an electrical charge, they cannot produce ionisation by electromagnetic interactions. Furthermore, slow neutrons don't have the energy to built such ionization by collisions. Thus, the only method to detect slow neutrons is by nuclear reactions, where the neutron is captured by the nucleus following secondary particles with high energies, which are able to ionize the detector material. So, neutrons need to be 'converted' to some secondary particles which are able to be detected.

The most important nuclear capture reactions for neutron detection are based on the following isotopes:

$$
\begin{array}{lcll}
^{10}B + n & \rightarrow & ^{7}Li + \alpha + 2.78\ MeV \quad (6\%) & \\
 & \rightarrow & ^{7}Li^{*} + \alpha + 2.31\ MeV \quad (94\%) & \rightarrow \quad ^{7}Li + \alpha + \gamma\ (0{,}48\ MeV) \\
^{6}Li + n & \rightarrow & ^{3}H + \alpha + 4.78\ MeV & \\
^{3}He + n & \rightarrow & ^{3}H + p + 0.764\ MeV & \\
^{155}Gd + n & \rightarrow & ^{156}Gd^{*}\ (8.54\ MeV) & \rightarrow \quad ^{156}Gd + \gamma\text{'s} + \text{ce's} \\
^{157}Gd + n & \rightarrow & ^{158}Gd^{*}\ (7.94\ MeV) & \rightarrow \quad ^{158}Gd + \gamma\text{'s} + \text{ce's}
\end{array}
$$

The capture cross sections of these isotopes for different energies are shown in Fig. 2. It is not only sufficient to offer a high cross section for neutron capture, but the secondary particles should also deposit sufficient energy for the detection. A large amount of ionization favours the discrimination against gamma-radiation.

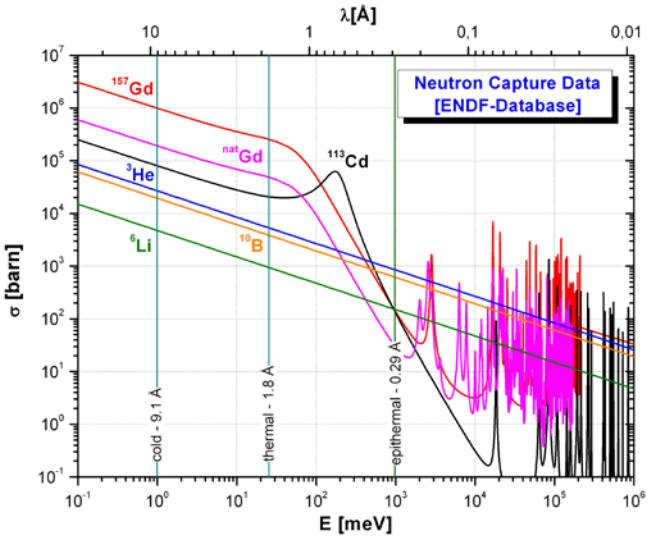


Fig. 2: Capture cross sections of the most important isotopes that are used as converter materials for slow neutrons. ^{113}Cd is also shown for comparison, because it is often used as shielding material.

3 Gaseous Detectors

The principle of operation for gaseous detectors is based on the ionization of gas by particle radiation within an electric field. The ions and electrons are moving to the electrodes by the electric field and induce a signal which can be detected by detector electronics. X-rays create this ionization by photoeffect, Compton-scattering or pair production. Neutrons create this ionization by nuclear reactions.

There are different operation regions for gaseous detectors in dependence of the applied electric field (voltage). The different regions are shown in Fig. 3.

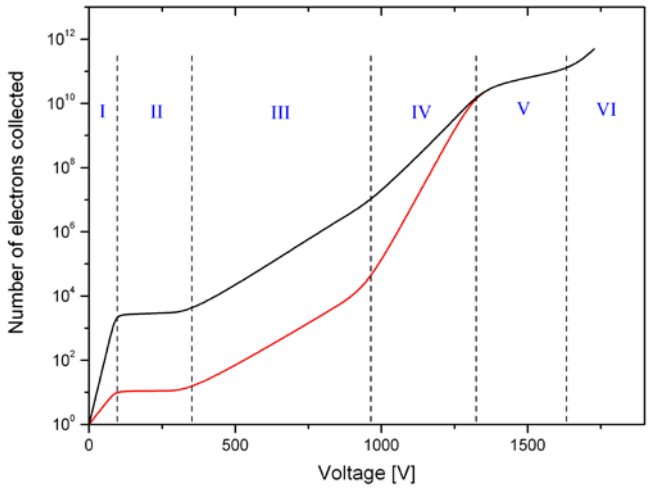


Fig. 3: Number of electrons collected per incident photon, illustrated for two different primary ionisations (e.g. two different photon energies)

Region I is called the recombination region. Here, the electric field is so low, that ion pairs partly recombine. The result is, that only a part of the charge carriers contribute to the pulse height of the detector. In region II, the ionisation region, all primary charge carriers are accumulated, within some range independent of the applied voltage. By rising the electric field in region III, further ionisation is produced by collision of the charge carriers. The gain of this amplification rises exponentially (avalanche), but the curves of different primary ionization are almost proportional to each other. In region IV the amplification gets more independent of the primary ionization. The proportionality between different primary ionizations is limited. Region V is called the Geiger-region. There is a number of charges collected, independent of the primary ionization. In region VI the electric field is so high, that discharges occur.

3.1 Ionization chamber

The ionization chamber is the simplest gaseous detector, which utilizes region II in the description above. An ionization chamber consists of a gas-filled parallel plate capacitor. In case of ionization by incident radiation, the electrons and ions are drifting separately to the anode and cathode. During this process a current is produced, which could be measured as a voltage signal.

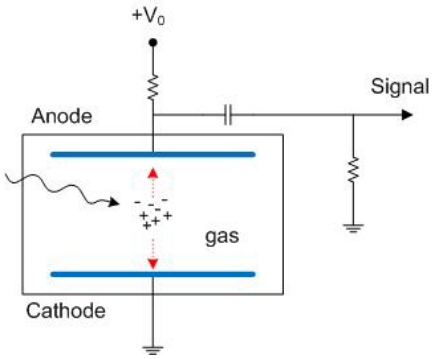


Fig. 4: Sketch of the operating principle of an ionization chamber.

The currents that need to be measured with ionization chambers are usually extremely small. If we assume that there are 1000 electron tracks per second in the gas volume and that each track is 2 cm long, the corresponding current is about 18 x 10^{-15} A. This ionization current causes a voltage drop over a resistor, which can be measured with a sensitive voltmeter (DC electrometer). The values of the resistors used are typically between 10^{9} Ω and 10^{12} Ω. Without adequate precautions by usage of guard rings the leakage current in the system could be much larger that the ionization currents one wants to measure.

Today, battery operated portable ionization chambers are commonly used as radiation monitoring instruments to measure gamma ray exposure. Then, the total charge pulse corresponding to one electron trajectory of a few centimeters long corresponds to only a few 100 electron charges. However, an ionization chamber can also be used to measure the concentration of air-borne alpha emitters (e.g. in smoke detectors). An alpha particle of 3.38 MeV will produce about 10^{5} electron–ion pairs in air. This is still a very small electrical pulse, but large enough to be observed with suitable low noise electronics.

3.2 Proportional Counter

Proportional counters work in the region with proportional charge multiplication (region II). Such detectors are made of a conductive cylinder with a thin wire in its centre. While the tube is at ground potential, a large positive voltage is applied to the anode wire.

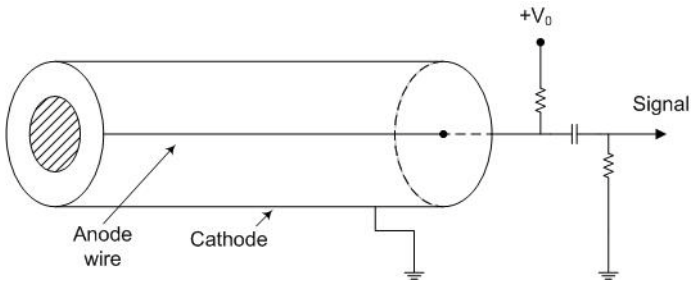


Fig. 5: Basic construction of a proportional counter tube.

The electric field for such a configuration shows a 1/r dependence. At large r the field is relatively weak, but becomes very high close to the surface of the wire. If ions and electrons are created in the cylindrical volume, they will simply drift towards their respective electrodes. But, if the electrons are very close to the wire, they are heavily accelerated by the high electric field, which leads to further ionization and finally to a charge avalanche. Regardless of where the ionizing event occurs, all multiplication takes place in a small region around the anode. The electrons of the avalanche are collected very quickly (~1ns) while the positive ions begin drifting towards the cathode. This ion drift is mainly responsible for the pulse signal on the anode wire.

The choice of filling gas is determined by several factors. High gain at low voltages, good proportionality and high rate capability are favorable properties. These conditions are generally achieved by gas mixtures rather than by pure gases.

Proportional counters are suitable for the detection of X-rays. The interaction probability of X-rays in argon drops to a very low value above 20 keV. With krypton or xenon, sizeable detection efficiencies up to100 keV can be obtained. For X-rays of the order of 10 keV, the photoelectron can be fully contained in the gas and the counter signal will be proportional to the energy of the X-ray. The energy resolution that can be obtained depends on the number of primary electron–ion pairs formed. However, the fluctuations on this number are not well described by Poisson statistics. The fact that the total energy used to create electron–ion pairs must equal the energy of the gamma reduces the fluctuations. The energy resolution of X-rays of 5.89 keV in argon is about 11% FWHM.

For neutron detection, either $^{10}BF_3$ or ^{3}He gas can be used. The incoming neutron reacts with the boron or helium isotope by a nuclear reaction to form the secondary reaction products. In case of ^{10}B a ^{7}Li and an α and in case of ^{3}He a ^{3}H and an α are produced and are moving in perpendicular direction from the reaction point through the gas. Thereby they lose their kinetic energy by collisions with electrons of the gas atoms, which leads to a creation of electrons and ions. This ionization is then detected by the process described above. As the reaction products have rather large path lengths in the gas, it is likely that one of secondary particles hit the wall of the tube, before it releases all its energy. This is seen in the pulse height spectra by a step-like form of the distribution.

Proportional tubes can also be used as position sensitive detector. By using a resistive wire, the ratio of charges at both ends of the wire determines the position. In neutron scattering, stacks of such detectors are used to cover large areas. The position resolution along the wire direction is typically about 1% of the wire length. The perpendicular position resolution is determined by the diameter of the tube.

3.3 Multiwire Proportional Chamber

A method to improve the position resolution of proportional tubes is the use of Multiwire Proportional Chambers (MWPC), developed by G. Charpak. In the simplest case a MWPC consists of gas volume with parallel anode wires between two cathode layers.

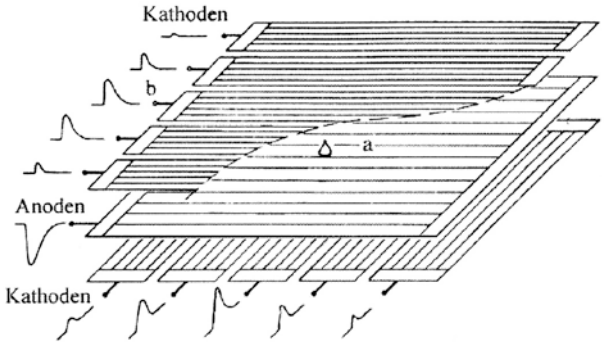


Fig. 6: Principle of the avalanche centre-of-gravity-method. The avalanches surrounding a wire induce a positive pulse on the cathode strips. The pulse height of the induced pulse is measured and the centre-of-gravity of the pulse height distribution is computed and gives the position of the avalanche.

With MWPCs a two dimensional position reconstruction is possible. The first coordinate is given by the position of the anode wire fired. The second coordinate is given either by cathode strips or by charge division of the anode wire. The position resolution is determined by the distance of the anode wires and the length of the ionization track. Typically, the position resolution is in the range of a few mm. The distance between the anode wires is determined by the Coulomb repulsion of the wires and is typically ~ 2mm. The local rate capability is about 10-20 kHz.

For X-ray detection the MWPC is filled with high-Z noble gases (e.g. Argon or Xenon) together with a quenching gas. For slow neutron detection ^{3}He is used together with CF_4 or Propane to reduce the track length. There, a high pressure is needed to achieve a sufficient detection efficiency which unfortunately also implies thick entrance windows. A spatial resolution of 0.4 mm has been achieved by a 8 bar ^{3}He and 6 bar propane gas mixture.

4 Semiconductor Detectors

A semiconductor detector, as its name implies, is a device that uses a semiconductor to detect particle radiation. Usually, doped silicon or germanium is used for semiconductor detectors, but today there are also other materials available, e.g. CdTe or CdZnTe, which can easily be used at room temperature.

In particle physics they are heavily used for detection of traversing charged particles. While semiconductor detectors play a minor role for neutron detection, they are very important for energy resolving X-ray detection.

4.1 General Operation Principle

The general operating principle of semiconductor detectors is based on the creation of electron-hole pairs by ionizing radiation, which are then collected by an electric field. Such an electric field is formed within the semiconductor detector by the method of a pn-junction with reverse bias voltage.

With the formation of a pn-junction a special zone is created at the interface between the two materials. The difference in the concentration of electrons and holes between the two materials leads to an initial diffusion of holes towards the n-region and a similar diffusion of electrons towards the p-region. As a consequence, the diffusing electrons fill up holes in the p-region while the diffusing holes capture electrons on the n-side. As the n and p regions were initially neutral, this recombination of electrons and holes causes a charge up on either side of the junction. This creates an electric field across the junction which counteracts and finally stops the diffusion process. A region of immobile space charge is then left, which is called the depletion zone. It has the special property that there are no mobile charge carriers. Any electron or hole created in this zone by ionizing radiation will be swept out by the electric field. As the average energy required to create an electron-hole pair is some 10 times smaller than in gaseous detectors, the amount of ionization produced for a given energy is an order of one magnitude larger, resulting in an increased energy resolution. Furthermore, the higher density also results in a higher stopping power.

4.2 Silicon p-i-n Detectors

Fig. 7 shows the basic configuration used for operating a junction diode as a radiation detector. Semiconductors cannot be contacted by directly depositing a metal on the semiconductor material, because such contacts usually result in the creation of a rectifying junction with a depletion zone extending into the semiconductor. In order to avoid this formation, heavily doped layers of n^+ and p^+ material are used between the semiconductor and the metal leads. Because of the high dopant concentrations, the depletion depth is then essentially zero, which forms the ohmic contact.

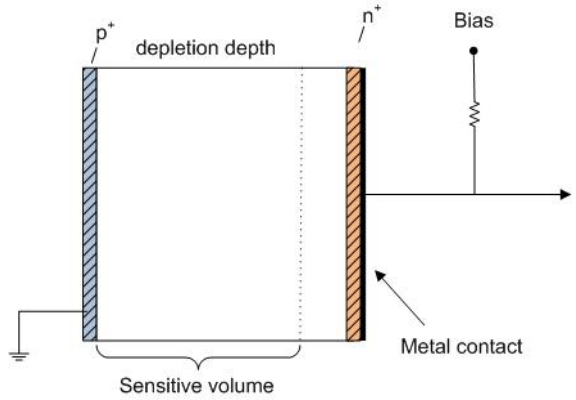


Fig. 7: The semiconductor detector is based on the diode with reverse bias voltage.

Simple silicon junction detectors have relatively small depletion depths, which are typically between 60 um and a few millimetres. They are particulary suitable for alpha and beta particles, but are relatively insensitive for X-rays because of the low absorption cross section. In order to obtain thicknesses larger than a few millimeters very high resitivities are required, which can only be obtained with intrinsic materials or eventually compensated semiconductor. This is achieved by the lithium drifting process for forming compensated material. Junctions formed with compensated materials are known as p-i-n junctions and posses properties different from regular pn-junctions. In particular, there is no space charge in the compensated zone, which implies an almost constant electric field.

Silicon p-i-n semiconductor detectors have large depletion depths of 10 mm and more. As the detection efficiency is a function of the thickness of the depletion layer, they can serve as efficient detectors for X-ray with a fairly good energy resolution. For example, for 5.9 keV X-rays from ^{55}Fe an energy resolution of 150 eV is routinely obtained with an detection efficiency of almost 100%. For higher energies above approximately 60 keV, photons interact almost entirely through Compton scattering. At these X-rays often fail to deposit all its energy in the silicon p-i-n detectors. Above 100 – 150 keV the detectors are seldom used for spectrometry.

4.3 Avalanche Diodes

With normal semiconductor detectors the objective is to collect all charge carriers that are created by the incident radiation. By raising the reverse bias voltage, a sufficiently high electric field is achieved, which enable the migrating electrons to create secondary ionization and finally an charge avalanche. The gain of such avalanche photodiodes is strongly dependant on the applied voltage and temperature. At stable operation conditions, gains of up to several hundred in the total collected charge are possible before breakdown.

The advantage of avalanche photodiodes exhibits if the photoelectric effect is used to convert light or soft X-rays to electricity. The charge multiplication permits use of avalanche photodiodes with sufficient signal-to-noise ratio, even for the detection of very low energy radiation. Thus they have found widespread application in the detection of low energy X-rays.

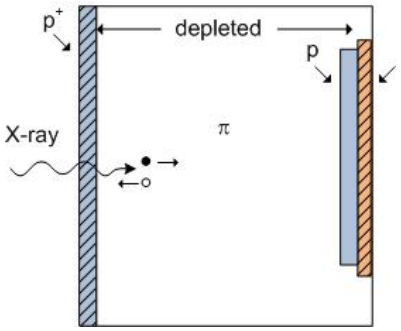


Fig. 8: Sketch of the avalanche photodiode operation principle. The incident radiation interacts in the π-region that constitutes most of the diode thickness. The electric field at the pn+ region is very high, so that charge multiplication occurs.

Silicon photomultipliers (SiPM) are silicon single photon sensitive devices built from an avalanche photodiode array on common Si substrate. The idea behind this device is the

detection of single photon events in sequentially connected Si APDs. The dimension of each single APD can vary from 20 to 100 micrometres, and their density can be up to 1000 per square millimeter. Every APD in SiPM operates in Geiger-mode and is coupled with the others by a polysilicon quenching resistor. Although the device works in digital/switching mode, the SiPM is an analog device because all the microcells are read in parallel making it possible to generate signals within a dynamic range from a single photon to 1000 photons for just a single square millimeter area device.

5 Scintillation detectors

Scintillation detectors make use of the fact that certain materials, when struck by nuclear radiation, emit a small flash of light, i.e. a scintillation. When coupled to a light detection device, e.g. a photomultiplier, these scintillations can be converted into electrical pulses which can be analysed and counted electronically to give information concerning the incident radiation.

5.1 Scintillation material

A scintillator can be described as a material that converts the energy of an incident particle or energetic photon into a number of photons of much lower energy in the visible or near visible range. There are both inorganic and organic materials showing this property, but the mechanism on which the scintillation is based is different for both types. While the emission of scintillation light in inorganic scintillators is based on the properties of the crystal lattice, in organic scintillators it is based on the molecular properties of the organic matter. Here, only inorganic scintillators are described, because this type is much higher importance for X-ray and slow neutron detection.

Inorganic scintillators are grown in a crystalline form and are doped with an activator material. Fig. 9 shows the electronic band structure of inorganic scintillators with its valence, exciton and conduction band. One criterium for the selection of the activator material is that its energy levels are located between the valence and exciton band.

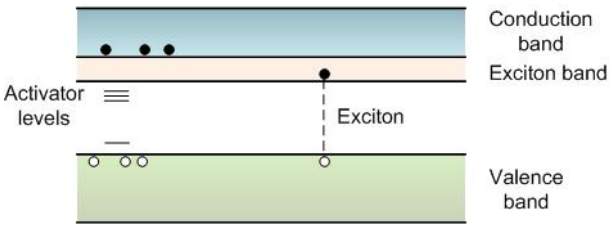


Fig. 9: Electronic band structure of inorganic scintillators. Besides the formation of free electrons and holes, loosely coupled electron-hole pairs known as excitons are formed. Excitons can migrate through the crystal and be captured by impurity centres.

When particle radiation enters the crystal, two principal processes can occur. It can ionize the crystal by an exciting an electron from valence band to the conduction band, creating a free electron and a free hole. Or it can create an exciton by exciting an electron to a band located just below the conduction band (exciton band). In this state the electron and hole remain bound together as a pair. However, the pair can move freely through the crystal. If the crystal now contains impurity atoms, electronic levels in the forbidden energy gap can be locally

created. A migrating free hole or a hole from an exciton pair which encounters an impurity centre, can ionize the impurity atom. If now a subsequent electron arrives, it can fall into the opening left by the hole and make a transition from an excited state to the ground state, emitting radiation if such a de-excitation mode is allowed. If the transition is radiationless the impurity centre becomes a trap and the energy is lost to other processes. Sometimes, the excited state is metastable, so the relaxation back out of the excited state is delayed (necessitating anywhere from a few microseconds to hours depending on the material): the process then corresponds to either one of two phenomena, depending on the type of transition and hence the wavelength of the emitted optical photon: delayed fluorescence or phosphorescence, also called after-glow.

Important properties of scintillators are the light yield, the emission wavelength and the decay time of the excitation. These properties largely depend on the activator material. The light yield is the amount of light delivered by the scintillator per incident particle energy. As the light is emitted by the activator atoms, the light yield is dependent on the concentration and the efficiency of the activator atoms in the crystal lattice. A high light yield is favourable, because it offers a better energy resolution. Also the emission wavelength and the decay time of the excitation are related to the activator atoms. The scintillation material should be transparent for the emitted light and the emission wavelength should be adapted to absorption characteristic of the light detection device. This is typically in the range of 400 nm, which corresponds to about 3 eV for the difference between ground and excited states of the activator. Finally, short decay times are favourable in order to achieve a high count rate capability of the scintillation detector.

Most commonly used activators for scintillators are Thallium, Europium or Cerium. There is a large variety of different scintillators available The most efficient converter of radiation energy to visible light is the thallium activated sodium iodide scintillator, NaI(Tl). It exhibits a light yield of about 40000 photons per MeV incident energy. Because of the relatively high atomic number of the iodine, it is well suited for gamma ray detection. However, the material is hygroscopic and must be encapsulated to protect it. The replacement of the sodium by caesium yields another very efficient scintillator, the CsI(Tl) scintillator. It is not hygroscopic and its stopping power for gammas is better than NaI(Tl). The light yield for a given gamma energy is 60000 photons per MeV and even higher as for NaI(Tl). But, the maximum of emission wavelength is around 540 nm, which is not well matched to the spectral sensitivity of photomultipliers. By usage of photomultipliers for light detection, the amount of light measured is only 60% of that for NaI(Tl). Standard scintillators for neutron detection are cerium activated glass scintillators, which contain a high amount of enriched lithium for the neutron capture. Such ^{6}Li-glass scintillators emit about 6600 photons per captured neutron with a peak wavelength of about 400 nm. It has a decay time of 75 ns, which allows for high count rates, and already 1 mm thickness is sufficient for an efficient slow neutron detection. Europium activated lithium iodide, LiI(Eu), is another important scintillator for slow neutron detection, which offers a higher light output than Li-glass. However, it is also highly hygroscopic and needs to be encapsulated. The brightest scintillator is silver activated zinc sulfide, ZnS(Ag). It is available in powder form, but it is not transparent. So, only thin layers of ZnS(Ag) scintillators can be applied. For slow neutron detection this material is mixed with enriched ^{6}LiF to form scintillators of thicknesses up to 450 µm.

5.2 Photomultiplier Tubes

A common method to detect light emitted by a scintillator is the use of photomultiplier tubes (photomultipliers or PMTs). Photomultipliers are vacuum tubes which are able to convert light into a measurable current and they are extremely sensitive in the ultraviolet, visible, and near-infrared ranges of the electromagnetic spectrum.

Fig. 10 shows a schematic diagram of a typical photomultiplier. A photomultiplier consists of a photosensitive cathode followed by an electron collection system, an electron multiplier section (dynodes) and finally an anode from which the final signal can be taken. The housing of a photomultiplier is an evacuated glass tube.

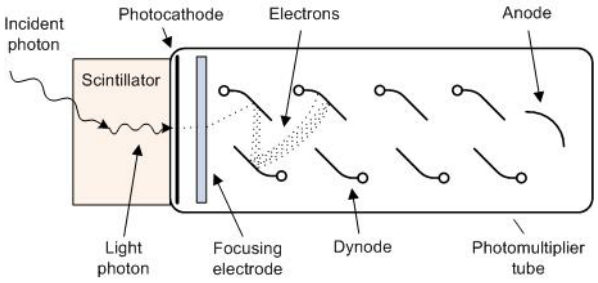


Fig. 10: Photomultiplier tube with head-on configuration.

During operation a high voltage is applied to the cathode, dynodes and anode such that a potential ladder is set up along the length of the cathode – dynode – anode structure. When an incident photon from a scintillator impinges upon the photocathode, an electron is emitted via the photoelectric effect. Because of the applied voltage, the electron is then directed and accelerated toward the first dynode, where upon striking, it transfers some of its energy to the electrons in the dynode. This causes secondary electrons to be emitted, which in turn, are accelerated towards the next dynode where more electrons are released and further accelerated. An electron cascade down the dynodes is thus created. At the anode, this cascade is collected to give a current which can be amplified and analyzed. If the cathode and dynode systems are assumed to be linear, the current at the output of the photomultiplier will be directly proportional to the number of incident photons. As the number of photons produced in the scintillator is proportional to the energy deposited, the output of the photomultiplier is also proportional to the energy that the particle radiation has left in the scintillator.

An important characteristic of photomultipliers is the quantum efficiency, which is the quotient of the number of photoelectrons released by the photocathode to the number of incident photons. This quantity is wavelength dependant and strongly related to the type of photocathode. Typically bialkali photocathodes are used which show a quantum efficiency of about 25% at a peak wavelength of 400 nm. The potentional between neighboring dynodes is adjusted by a voltage divider circuit and is in the range between 50-200 V. Photomultipliers are usually equipped with 10-16 dynodes. With a secondary emission factor of the dynodes of 3-4, this results in overall amplification of 10^5 - 10^9.

Photomultipliers are rugged light detection devices with a high amplification. Nevertheless photomultipliers are also influenced by environmental factors. The most critical factor are external magnetic fields. It is obvious, that already small magnetic fields are sufficient to

deviate the electron cascade from its optimum trajectory and thereby affect the operation characteristic. For small magnetic fields a shielding with a mu-metal screen around the photomultiplier may help, but at higher magnetic fields, other light detection devices, e.g. avalanche photodiodes, must be used.

5.3 Anger Camera Scintillation Detectors

The Anger camera principle is a simple well proven method to build cost efficient large area, position sensitive scintillation detectors. With this method, the light emitted by an impinging particle in the scintillator, is spread via a light disperser over several photomultiplier tubes. By analyzing the signals of the photomultiplier tubes, the event can be computed (e.g. by centre-of-gravity calculation). Thus, with a small number of photomultiplier tubes, a spatial resolution may be achieved, which is much smaller than the tube dimensions.

Anger camera detectors have been built for X-ray and neutron detectors. A construction for slow neutron detection is shown in Fig. 11, which utilizes a thin ^{6}Li-glass scintillator for neutron capture. In practice, there is a small air-gap between the scintillator and the light disperser. Thus, by total reflection in the scintillator, the light spread is limited to a cone that strikes the photomultiplier array. This is mainly to facilitate the reconstruction method. For Anger camera detectors, the thickness of the light disperser has to be adapted to the diameter of the photomultipliers.

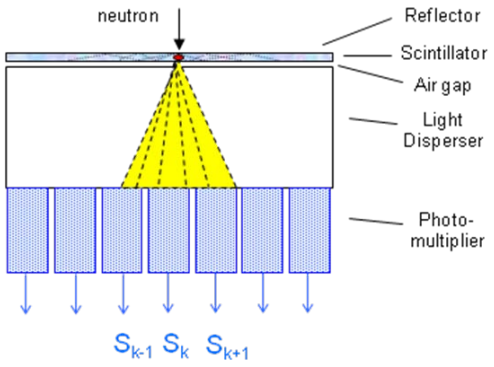


Fig. 11: Sketch of the Anger camera principle. Light produced by the neutron capture in the scintillator is spread over an array of photomultipliers. By analysing the pulse heights of the photomultiplier, the position of the neutron capture can be computed.

6 Integrating Detectors

As the name indicates, integrating detectors are not readout on a event basis, but rather accumulate information about radiation interactions over a longer period of time. The simplest case of such detectors is the photographic film used for the detection of X-rays in medical imaging. An advantage of integrating detectors is usually the high position resolution, but due to the missing event information interfering effects cannot be removed.

6.1 Image Plates

Image Plates are integrating position sensitive detectors that are based on photostimulable phosphors. By exposure of such materials with ionizing radiation, the excited electrons in the

phosphor material remain trapped in colour centres of the crystal lattice until stimulated by the second illumination.

The commonly used material is BaFBr:Eu^{2+}, in which the ionizing radiation partly ionizes Eu^{2+} ions to Eu^{3+}. The lose electron enters the valence band of the crystal and becomes trapped in the ion empty lattice of the crystal. By illumination with visible light (e.g. He-Ne-laser) the trapped electrons are lifted to the valence band and are catched by Eu3+ ions. Thereby characteristic Eu2+ luminescense light is emitted, which can be detected by photomultipliers. The intensity is proportional to the radiation deposited. After readout of the image plate the whole information is erased by intense white light. Then it can be used for new images.

Image plates can directly be used for X-ray detection. In case of neutrons the material is made sensitive by mixing it with Gd_2O_3-powder. Neutrons are then absorbed by ^{155}Gd or ^{157}Gd and registered by the subsequent emission of conversion electrons, Auger electrons and X-rays. The detection efficiency for thermal neutrons is small (~23%) and because of the high-Z materials it is gamma sensitive.

6.2 Charge Coupled Device (CCD)

A charge-coupled device (CCD) is a silicon device with a two dimensional array of tiny potential minima, each covering an area of a few microns. When struck by radiation, electrons are released which are then trapped in these minima. This charge information is then readout by successively shifting the charge from one minimum to the next until it reaches the output electronics. CCD's are mainly used for light imaging purposes and are extremely sensitive, low noise devices. It can be used as a large X-ray detector in combination with a phosphor screen and a fiber optic taper as shown in Fig. 12.

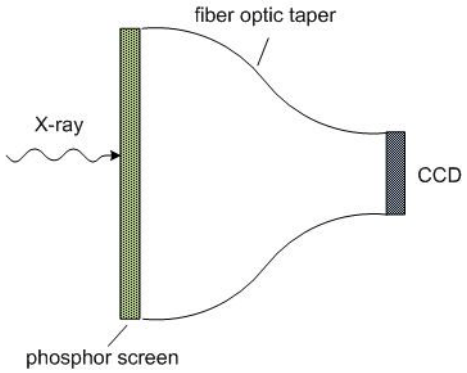


Fig. 12: Building blocks of a X-ray CCD detector: X-rays are first converted by the phosphor screen. The fiber optic taper focus the image on the CCD, where the photons are detected.

As phosphor screen most commonly $Gd_2O_2S(Tb)$ is used, which emits mainly green light. This material is available as powder which is bound with a binder material onto an X-ray transparent window (aluminized Mylar). The fiber optic taper is to focus the output image of the phosphor screen onto the CCD chip. It usually consists of many glass fiber light guides, which are melted and stretched to create a typical demagnification. Depending on this demagnification, there is a large attenuation of the number of photons by the fiber optic taper. Typically, only about 20% of the emitted photons are finally detected by the CCD.

References

[1] Knoll G., Radiation Detection and Measurement (Wiley, New York, 2000)
[2] Leo W., Techniques for Nuclear and Particle Physics Experiments (Springer, 1987)
[3] Tavernier S., Experimental Techniques in Nuclear and Particle Physics (Springer, 2010)
[4] Lecoq P., Inorganic Scintillators for Detector Systems (Springer, 2006)
[5] Lutz G., Semiconductor Radiation Detectors (Springer, 1999)

C 6 Polarized Neutron Scattering and Polarization Analysis [1]

W. Schweika

Jülich Centre for Neutron Science

Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Among the properties that make the neutron a unique and valuable probe for condensed matter research, its spin and magnetic moment is of particular importance in the scattering process. There are two aspects to consider, firstly, the strong nuclear interaction of the neutron with an nucleus depends on the either parallel or antiparallel alignment of the spins of neutron and nucleus, and secondly, the neutron's magnetic moment interacts with the unfilled electron shells of atoms or ions in magnetic scattering. The scattering process will also have an impact on the spin state of the neutron probe. Hence, one may expect that controlling the neutron spin in a scattering experiment will provide further valuable information, which leads us to the subject of this lecture: polarized neutron scattering and polarization analysis. Indeed, it is possible by working with polarized neutron and polarization analysis to separate scattering terms with respect to their different structural or magnetic origins, and moreover, to uncover scattering contributions that remain hidden in usual unpolarized scattering experiments.

2 Neutron spins in magnetic fields

The neutron has a spin $S = \pm 1/2$ with angular momentum $\mathbf{L} = \hbar\mathbf{S}$. The magnetic moment of the neutron results from the spins of the individual quarks and their orbital motions, and the relation between spin and magnetic moment is given by the neutron g-factor, $g_n = -3.8260837(18)$, in units of the nuclear magneton $\mu_N = \frac{e\hbar}{2m_p} = 5.05078324(13) \cdot 10^{-27}\mathrm{JT}^{-1}$

$$\mu_n = g_n S \mu_N \simeq \mp 1.913 \mu_N = \pm \gamma_n \mu_N,$$

where $\gamma_n = -g_n S$ is the gyromagnetic factor of the neutron (see Refs. [1, 2]) . Because of the small ratio $\mu_n/\mu_B = m_e/m_p$, the neutron magnetic moment μ_n is small compared to the magnetic moment of the electron $\mu_e = g_e S \mu_B \approx 1\mu_B$, with the Bohr magneton $\mu_B = \frac{e\hbar}{2m_e}$ and the Lande-factor $g_e = 2(1 + \alpha/2\pi - 0.328\alpha^2/\pi^2) \approx 2$ (see Refs. [3, 4]), α is the fine structure constant. A peculiarity to note is that different to the electron and proton, the neutron magnetic moment is aligned opposite to its spin.

2.1 Interaction of neutrons with magnetic fields

The dipolar interaction potential of a neutron with the magnetic field is given by $V_M = -\boldsymbol{\mu_n} \cdot \mathbf{B}$, where $\mathbf{B}$ is the magnetic induction. For neutrons passing from zero-field into a magnetic field the potential energy changes by the Zeeman term $\pm\mu_n \mathrm{B}$ depending on the relative orientation of the magnetic moment. The according change in kinetic energy to conserve the total energy is small, $0.06\,\mu\mathrm{eV/T}$, which in experimental practice is of relevance only in rare cases [5].

Larmor precession

The characteristic motion of the neutron magnetic moment in a magnetic field is Larmor precession, which for simplicity can be considered in a classical treatment. In fact, even the quantum mechanical treatment, which introduces Pauli spin matrices $\hat{\boldsymbol{\sigma}}$ into the Schrödinger equation, is effectively a classical treatment considering the origin of these matrices. Originally[6], they result from the problem of mapping three dimensions onto two by introducing a complex component describing the classical problem of a spinning top.[7]

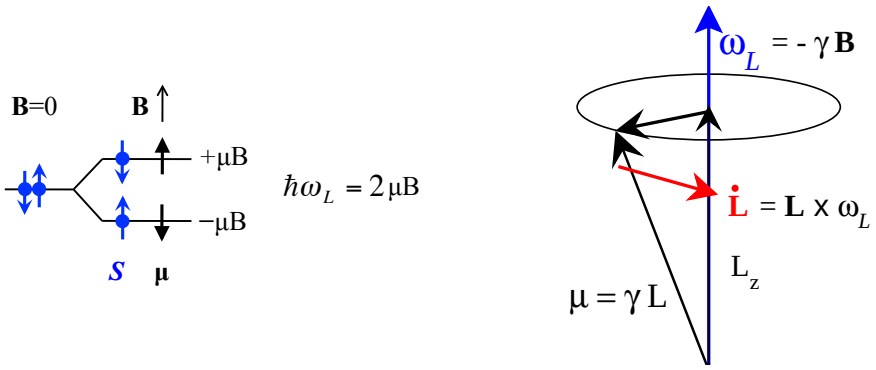


Fig. 1: *Larmor precession: the motion of the neutron in a constant magnetic field conserves energy and angular momentum.*

The magnetic interaction tends to align the neutron moment with the magnetic induction in order to minimize the interaction energy. The magnetic moment is related to the angular momentum as $\boldsymbol{\mu} = \gamma \mathbf{L}$, where γ is the gyromagnetic ratio. Therefore, there is a torque $\boldsymbol{\mu} \times \mathbf{B} = \dot{\mathbf{L}}$ equal to the time derivative of the angular momentum, which leads to the Bloch equation of motion:

$$\dot{\boldsymbol{\mu}} = \gamma \, \boldsymbol{\mu} \times \mathbf{B}. \tag{1}$$

The gyromagnetic *ratio* for the neutron, not to be confused with the gyromagnetic factor, is given by $\gamma = 2\gamma_n \mu_N / \hbar = -1.83 \cdot 10^8 \, \mathrm{s}^{-1}\mathrm{T}^{-1}$ or, in cgs units, $\gamma/2\pi = -2916 \, \mathrm{Hz/Oe}$. Because of the cross product, the time derivative of the magnetic moment is always perpendicular to the moment itself. Therefore, the resulting motion is a precession, where the angular momentum, the component L_z along the field, and the energy are constants of the motion, see Fig.1. The precession frequency is the *Larmor frequency* $\omega_L = -\gamma B$, and $\hbar\omega = 2\mu B$, the Zeeman splitting energy.
Next we consider the many particle ensemble of neutrons in a neutron beam. Polarization of a neutron beam is defined by the normalized average over the neutron spins.

$$\mathbf{P} = 2\langle \mathbf{S} \rangle \tag{2}$$

In an applied field the individual neutrons split in spin up $n_\uparrow$ and and spin down states $n_\downarrow$. Since polarization will be measured with respect to a magnetic field defining a quantization axis, any device for polarization analysis will take the projection of the spins in up- and down state states,

$$\mathrm{P} = \frac{n_\uparrow - n_\downarrow}{n_\uparrow + n_\downarrow}. \tag{3}$$

Motion in time dependent fields

The polarization will behave like the individual neutron spin in a constant magnetic field and will be a constant of motion. However, if we consider time-dependent fields, the finite velocity distribution in a neutron beam will result in a different time dependence. Hence dephasing of neutrons spins and degrading of the polarization are possible experimental effects and have to be taken into account.
Thermal neutrons move with a speed of thousands of meters per second. When passing through spatially varying magnetic fields, the neutrons experience time-dependent field changes in their

reference system. Replacing the constant $\mathbf{B}$ by $\mathbf{B}(t)$, the differential equation Eq. (1) can be used to calculate numerically the effect of all relevant field variations in an experimental set-up. Usually, it is possible to work within two simple limiting cases of either **(i) slow adiabatic field variation**, in which the non-precessing spin component parallel to the field smoothly follows the field direction, or of **(ii) sudden field reversal**, in which the non-precessing spin component has no time to reorient itself, when traversing abruptly from a parallel to anti-parallel field or vice versa. Slow field variation means that the field $\mathbf{B}$ changes or rotates in the coordinate system of the neutron with a frequency that is small compared to the Larmor frequency, see Fig. 1.

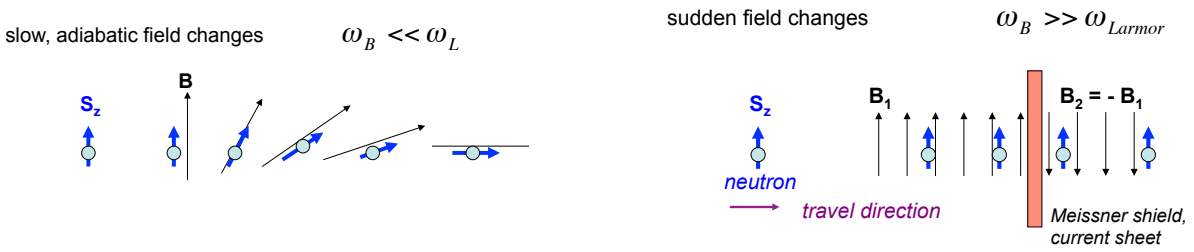


Fig. 2: *Neutron polarization can be best preserved in the asymptotic cases of either slowly or suddenly varying fields. The second case is used in a cryoflipper to reverse the polarization with respect to the external field.*

2.2 Experimental devices

Polarizer and polarization analyzer

The most common methods to polarize neutrons are (i) using the total reflection from magnetic multi-layers, (ii) using Bragg reflection of polarizing single crystals (typically Heusler crystals) and (iii) polarized He-3 filters, in which for anti-parallel spins the (n,^{3}He)-compound has a large absorption cross-section while all neutrons with parallel spins may pass the filter cell. The first two methods use an interference effect of nuclear and magnetic scattering amplitudes having the same absolute value.

(i) Polarizing total reflecting supermirrors are an easy experimental tool to obtain a broad wavelength band of cold polarized neutrons. The angle of total reflection for a single ferromagnetic (FM) layer is given by

$$\Theta_c^{\pm} = \lambda\sqrt{n(b-p)/\pi}. \tag{4}$$

Here n denotes the particle density and b and p the nuclear and magnetic scattering lengths, respectively. However, the critical angle can be further increased by artificial multi-layers (supermirrors) of alternating FM and non-magnetic layers of varying thickness [8], see Fig. 3. The λ dependence of the total reflectivity makes this type of polarizer less favorable for thermal neutrons of shorter wave length as it reduces the accepted divergence of the beam.

(ii) Bragg reflection by Heusler crystals is alternatively used to polarize thermal neutrons. Cu_2MnAl is a Heusler alloy, its (111) Bragg reflection gives 95 % polarization. However, the reflectivity of such crystals is low compared to usual non-polarizing monochromators and a saturating field is required over the entire monochromator, which makes it technically more complicated to combine with focusing though this is feasible.

(iii) He-3 filter cells, see Fig. 4, are of growing importance for polarizing neutrons, particularly for the more challenging case of thermal neutrons, although such cells are technically rather demanding and still under development. In case of spin-exchange optical pumping (SEOP) the spin polarization of ^{3}He gas is achieved in several steps. The cell is filled with additional Rb, K and N_2 vapor. Rb electrons are polarized with a a circular polarized laser, by collisions the spin is exchanged with K, which most efficiently exchanges spin with ^{3}He. Since polarization results from absorption, such a device does not interfere with the beam divergence. One may note that requirements for field homogeneity are very high and it is a kind of art and glass alchemy to make cells with small depolarization all determining the lifetime of He-3 polarization. The neutron polarization P raises with increasing He-3 cell size or pressure, while the transmission T decreases. The optimum in efficiency is usually chosen by the maximum of a quality factor P^2T.

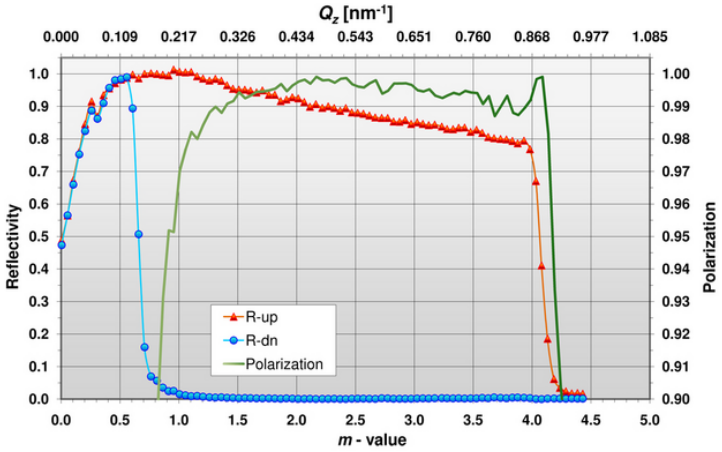


Fig. 3: *Spin dependent reflectivity and polarisation of a Fe/Si polarizing supermirror. The degree of polarization is high, note the expanded scale* . m=1 *($Q = 0.0217$Å^{-1}) corresponds to the total reflectivity of natural Ni. (from [9])*

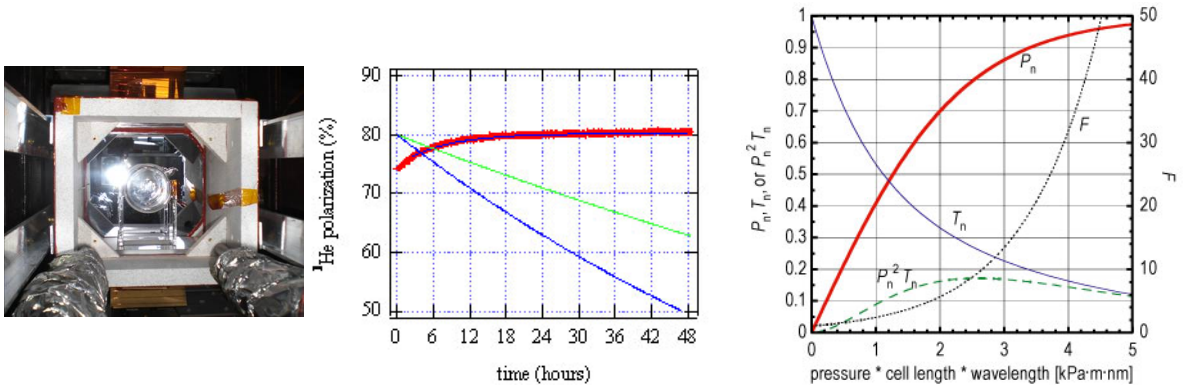


Fig. 4: *In-situ polarization of a SEOP He-3 cell and measured polarization of neutron beam in transmission. (from Babcock et al.[10])*

Guide-fields

A magnetic *guide field* is used to maintain the direction of the spin and the polarization of the neutron beam. The guide field preserves the quantization axis to which the neutron moments have align either parallel or anti-parallel. Typical guide fields in the order of 10 G are strong compared to earth field and other possible stray fields and usually sufficient to prevent depolarization along the beam path. Such guide fields are usually also too weak to have any significant impact on the sample magnetization.
Depolarization effects may for an inhomogeneous distribution of field directions over neutron beam cross-section, which is typically a few cm^2. This can easily occur if ferromagnetic material is used close to the beam or if the sample itself is ferromagnetic. A neutron beam passing through a ferromagnet is usually completely depolarized by differently oriented ferromagnetic domains in the beam path, unless a saturating magnetic field is used to align the domains. Depolarization will depend also on the path length through the sample, therefore, usually such effects are negligible in *neutron reflectometry* of thin ferromagnetic films.

Flipper

The purpose of a **π-flipper** is to reverse the polarization and to detect whether the sample causes spin-flip scattering.
When applying a magnetic field perpendicular to the polarized neutron beam, the polarization immediately starts its Larmor precession. A flipper that reverses the neutron polarization with respect to the guide field has to induce a well-defined field pulse so that the polarization precesses by an angle π. For this purpose one can use the homogeneous field of a $Mezei$ flipper, a long rectangular coil, see Fig. 5. Neutrons see a sudden field change when they enter and exit the coil, in between they precess around the perpendicular flipping field, whose magnitude is tuned with respect to the time of flight that the neutrons spend inside the coil. Thus, for a π flip in coil of thickness d, the time-of-flight $t = d\lambda\frac{m}{h} = \pi/\omega = -\pi/\gamma B$, (for example λ=4Å, B=17G, d=1cm). Just to mention, other types of flippers exist like radio frequency (RF) flipper (setup Moon, Riste Koehler, Fig. 8 and cryoflipper (Fig.2 and CryoPad Fig.10).

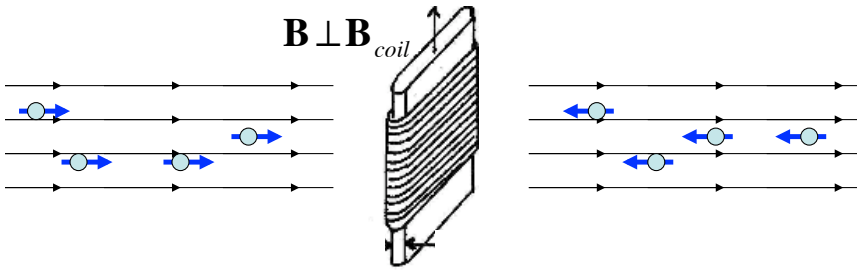


Fig. 5: *Principle of a neutron π-spin flipper. The neutrons perform a Larmor-precession of* 180^o *inside a long rectangular coil. The field $\boldsymbol{B}$ is perpendicular to spin orientation and adjusted to the speed of the neutrons.*

The purpose of a **π/2-flipper** is to set the polarization in precession mode by turning the polarization perpendicular to the guide field. A precessing polarization is used for instance in high resolution *Neutron Spin Echo spectroscopy* and for *Larmor diffraction*, see below. Both methods use the property of the neutron spin as an internal clock independent of the scattering process itself and achieve highest resolution.

Precessing polarization can also be used for analyzing polarization dependent scattering. A polarization analyzer accepts neutrons with spins aligned with respect to the magnetic guide field axis. If prior to the scattering process the neutron polarization is precessing with a known phase and if then finally we detect a neutron polarization along the guide field, we can conclude that scattering has caused the observed rotation of polarization. This method can be used for *neutron polarimetry* or *spherical polarization analysis*.

XYZ-coils

In order to align the polarization to any desired direction at the sample position, there is in the simplest version a set-up of three orthogonal pairs of so-called xyz-coils. Fig. 6 illustrates the field setting along x-direction. With such a device one can probe the direction of magnetic moments in the sample. One can see that the z-coil has been used to compensate the guide field at the sample position, and that the x-coils produce a field of a few Gauss. The field needs to be sufficiently strong so that the neutron polarization can follow the smooth variation of the field direction adiabatically, finally turning back into the z-direction of the guide field outside the xyz-coils.

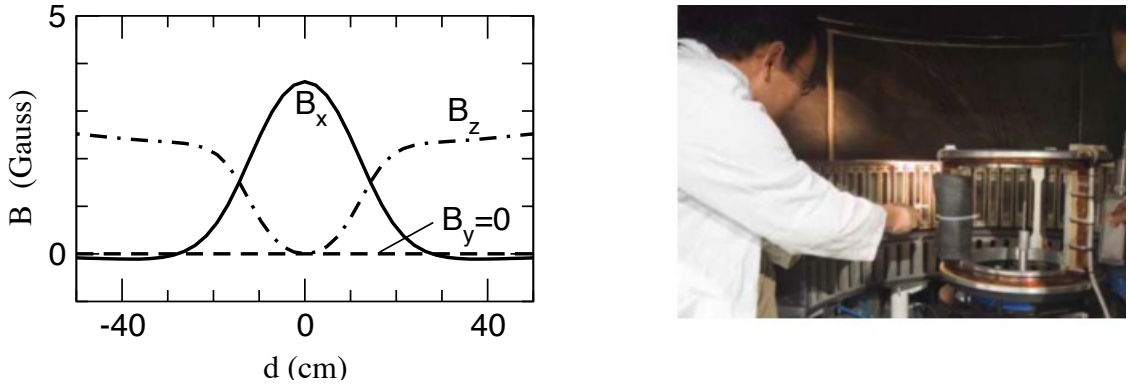


Fig. 6: *(left) Magnetic field setting in a xyz-coil system for an adiabatic nutation of the polarization of cold neutrons in horizontal x-direction at the sample turning to a vertical (guide) field B_z at further distance from the sample. (right) A photo of the xyz-coil system in the DNS instrument at the FRM-2.*

2.3 Polarized neutron instrumentation

The spin of the neutron and its Larmor precession in a magnetic field is an individual property of each neutron that can also used to measure indirectly the momentum or energy transfer in a scattering process, and with extremely high resolution. The ideas of the two most interesting applications of neutron precession spin-echo spectroscopy and Larmor diffraction will be discussed next.

Of course, the main objective of polarized neutron work is determine the polarization dependence of scattering. The selected examples below are important but cannot give an exhaustive overview on the suite of polarized neutron instrumentation, which is of growing diversity.

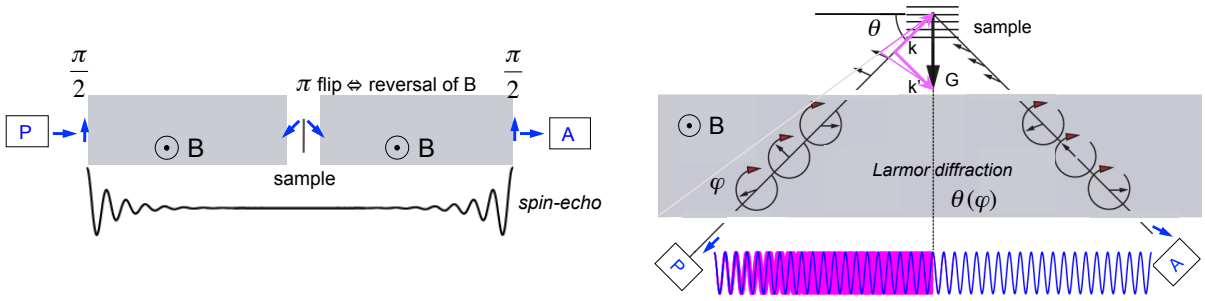


Fig. 7: *Precession in spin-echo spectroscopy (*left*) and Larmor diffraction (*right*, adapted from [13]).*

Neutron precession techniques

The idea of **neutron spin echo spectroscopy** is to use the individual spin precession of the neutrons to monitor possible small energy transfers of approximately 1μeV upon scattering. The incident beam is polarized first parallel to the field, then set into precession mode by a $\pi/2$ flipper and along the path through the precession coil, the total precession angle of the neutron is increases proportional to the time in the coil. Because $t \propto \lambda$, soon neutron spins are out of phase and the polarization vanishes to zero, which however can be fully recovered. Therefore, in the simplest case of elastic scattering, an equivalent field of opposite direction in the secondary flight path will turn the precession angles of all neutrons backwards and a spin echo - after another $\pi/2$ flip - can be detected by the final polarization analyzer. Energy exchange with the sample changes the average neutron speed and to recover the full spin echo the field integral has to be varied and adjusted. For a more precise understanding of the general inelastic case, we note that this is a special diffraction experiment, in which the energy transfer integral of the scattering law $S(\mathbf{Q},\omega)$ is weighted with $\mathbf{P}$ $(= P_z)$:

$$I \propto \int d\omega\, P_{(z)} S(\mathbf{Q},\omega) = \frac{1}{2}(S(Q) + \int d\omega\, \cos(\phi) S(\mathbf{Q},\omega)),$$

where $S(\mathbf{Q})$ gives the average from the unpolarized case. With $\phi = \omega t = \text{const} \cdot \omega \int ds\, B$ and identifying $\text{const} \int ds\, B = t$ as a Fourier time, one obtains

$$I \propto \frac{1}{2}(S(Q) + S(Q,t)).$$

In **Larmor diffraction**, all of the neutrons, which make the same Bragg reflection undergo the same number of precessions. Hence, the full neutron polarization is preserved and can be measured; there is no dephasing because of wavelength spread, beam divergence and mosaic of the sample. The total precession angle ϕ is proportional to the time t in the magnetic field region, and $t = s/v =$ is constant for a given Bragg reflection, since $s(\theta) \propto k$ and $v \propto k$, any spread in k distribution is compensated. Furthermore, the sum of incoming and outgoing flight path through the precessing region remains practically constant for a mosaic of the crystal, *i.e.* small tilts of the Bragg planes. Moreover, a combination of Larmor diffraction with the spin-echo principle, by introducing a π-flip near to the sample, allows to measure the mosaic by a loss in polarization.[11] Absolute d-spacings have been measured with $7 \cdot 10^{-5}$ Å accuracy.[12]

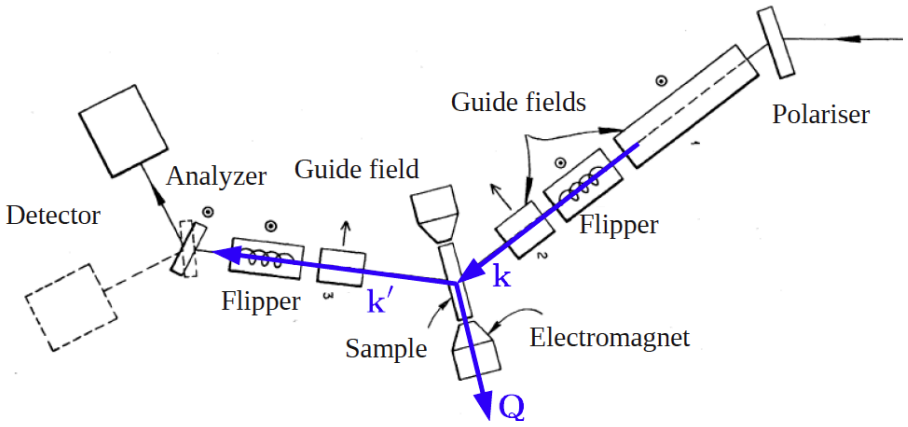


Fig. 8: *Setup for polarization analysis on a triple axis instrument by Moon* et al.*(1969)[14], see also text.*

Uniaxial polarization analysis on 2- and 3-axis instruments

Uniaxial polarization analysis means that not the full final polarization but only the projection of the final polarization with respect to an applied field is measured. There are a large number of triple axis instruments (TAS), which offer in various technical realizations the principle of uniaxial polarization analysis. ln the original setup by Moon, Riste and Koehler[14] at the Oak Ridge reactor (HFIR), see Fig.15, Co-Fe crystals mounted in the gap of permanent magnets are used on the first and third axis for the production of the polarized monochromized beam and for analysis of the scattered neutrons in energy and spin state. At the second axis with the sample an electromagnet can be turned to set the field either vertically or horizontally. Radio-frequency coils are used as flipping devices.

XYZ polarization analysis on time-of-flight multi detector instruments

The diffuse neutron scattering spectrometer DNS at the FRM2 in Munich is equipped with polarization analysis and is particularly devoted to elastic and inelastic diffuse scattering that may arise from spin correlations and magnetic disorder and ordering in materials. A layout of the instrument is shown in Fig. 9. DNS is a time-of-flight instrument[15] with a multi-detector system similar to the D7 instrument at the ILL [16, 17]. The monochromatic incident beam is polarized with a focusing supermirror bender, xyz-field coils allow for a change of the polarization at the sample, and the polarization analysis is performed with supermirror analyzers in focusing arrangement in front of each detector.

Zero-field spherical polarization analysis - CryoPad

The CryoPad is a portable, cryogenic device developed at the ILL by Tasset and LeLievre-Berna, which avoids any magnetic field and precession at the sample and achieves full and accurate control of change and rotation of incident to final polarization. As scheme and a set-up on the POLI-HEIDI instrument at FRM-2 is depicted in Fig.10.

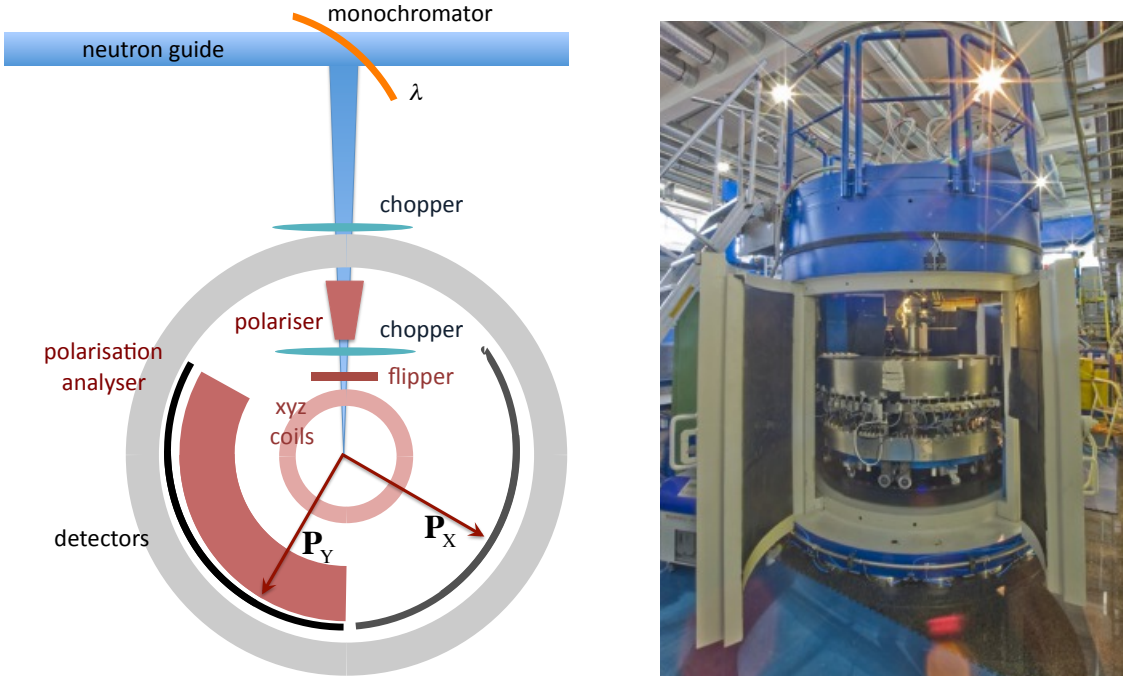


Fig. 9: *Polarization analysis on a time-of-flight multi-detector instrument, the DNS instrument at FRM-2, see Ref.[30] for the specific setting of polarization.*

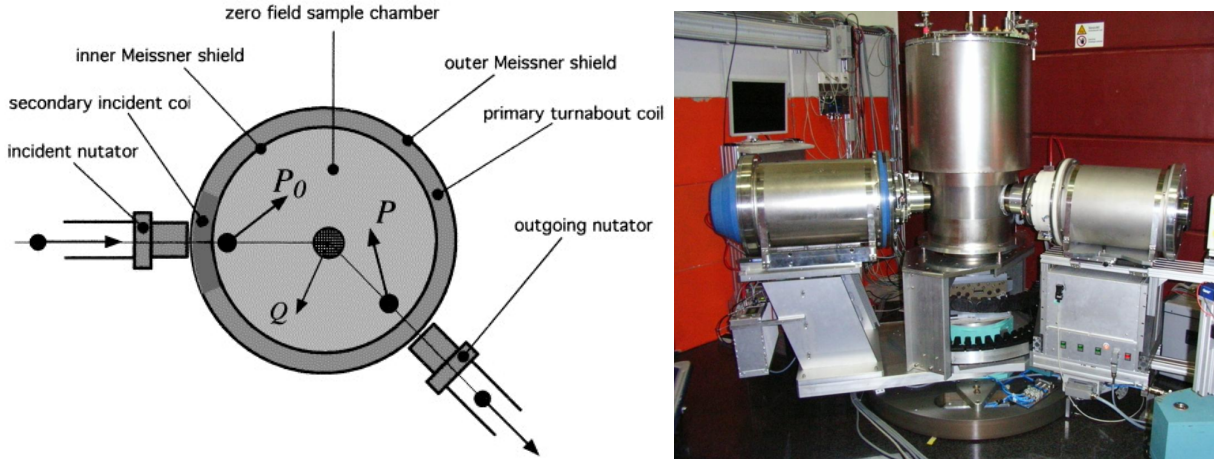


Fig. 10: *CryoPad (scheme)[18] and set-up at POLI-HEIDI (FRM-2). The zero-field around the sample avoids a precession of the beam polarization and allows to measure changes in magnitude and direction of* $\mathbf{P}$ *by scattering.*

3 Scattering and polarization

3.1 Interaction of neutrons with matter

The neutron scattering amplitude $F_{\mathbf{Q}}$ is determined by the transition matrix elements for a given scattering potential $V_{\mathbf{Q}}$

$$F_{\mathbf{Q}} = \langle \mathbf{k}'\mathbf{S}'|V_{\mathbf{Q}}|\mathbf{k}\mathbf{S}\rangle \tag{5}$$

resulting in a scattering cross section that in general depends on the scattering vector $\mathbf{k}-\mathbf{k}' = \mathbf{Q}$ and also on the energy transfer, which for simplicity will not be included in the notation:

$$\frac{d\sigma}{d\Omega} = \left(\frac{m_n}{2\pi\hbar^2}\right)^2 |F|^2 \tag{6}$$

The **nuclear interaction** operator $\hat{V} = (2\pi\hbar^2/m_n)\hat{b}$ can be described by a point-like and isotope-specific Fermi potential. For nuclei with zero spin (*e.g.* ^{12}C, ^{16}O ..., and typically *"gg"* isotopes with even number of protons and neutrons) the scattering length operator $\hat{b}$ is a scalar and the scattering will be independent of the neutron spin orientation.
The interaction is spin-dependent if the scattering nuclei have a non-zero spin $\mathbf{I}$ and the scattering lengths differ for parallel and antiparallel alignment of $\mathbf{I}$ and $\mathbf{S}$, which can be written as the sum of an average and coherent part A and a fluctuating spin-dependent part

$$\hat{\mathrm{b}} = A + B\ \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{I}} \tag{7}$$

where $\hat{\boldsymbol{\sigma}}$ is the Pauli spin operator given by Pauli spin matrices $\hat{\sigma}_{\mathrm{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$, $\hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Defining a quantization axis z for the neutron polarization $\mathbf{P} = 2\langle\hat{\mathbf{S}}\rangle = \langle\,\hat{\boldsymbol{\sigma}}\rangle$, with spin-up states $|+\rangle = \binom{1}{0}$ and spin-down states $|-\rangle = \binom{0}{1}$ we obtain the transition matrix elements

$$\langle +|\ \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{I}}|+\rangle \;=\; I_z, \tag{8}$$

$$\langle -|\ \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{I}}|+\rangle \;=\; I_x + \mathrm{i}I_y, \tag{9}$$

for non-spinflip and spinflip scattering amplitude, respectively. Therefore, two thirds of the spin-incoherent scattering is spinflip scattering. The final polarisation is given by $\mathbf{P}' = -\frac{1}{3}\mathbf{P}$. There is a change in sign and a reduction in magnitude, but no inclination towards $\mathbf{P}$. We may note that in contrast to the dipolar magnetic interaction, as discussed below, the obtained result is independent of the direction of $\mathbf{P}$ with respect to $\mathbf{Q}$.

In summary, we can distinguish three contributions to the nuclear scattering $|N_{\mathbf{Q}}|^2$ arising from the total nuclear scattering amplitude $N_{\mathbf{Q}} = \sum_j b_j e^{i\mathbf{Q}\cdot\mathbf{R}_j}$, the average coherent scattering, the isotopic, non-spin dependent part of the incoherent scattering, and the spin-incoherent scattering

$$\frac{d\sigma}{d\Omega}^N_{\mathbf{Q}} = |N_{\mathbf{Q}}|^2 = \frac{d\sigma}{d\Omega}^N_{\mathbf{Q},coh} + \frac{d\sigma}{d\Omega}^N_{isotop-inc} + \frac{d\sigma}{d\Omega}^N_{spin-inc}. \tag{10}$$

In absence of magnetic scattering, the sum of the coherent and isotopic incoherent nuclear scattering can be separated from the spin-incoherent scattering by measuring spin-flip and non-spin-flip scattering.

$$\frac{d\sigma}{d\Omega}^N_{\mathbf{Q},coh} + \frac{d\sigma}{d\Omega}^N_{isotop-inc} = \frac{d\sigma}{d\Omega}^{NSF} - \frac{1}{2}\frac{d\sigma}{d\Omega}^{SF} \tag{11}$$

$$\frac{d\sigma}{d\Omega}_{spin-inc} = \frac{3}{2}\frac{d\sigma}{d\Omega}^{SF} \tag{12}$$

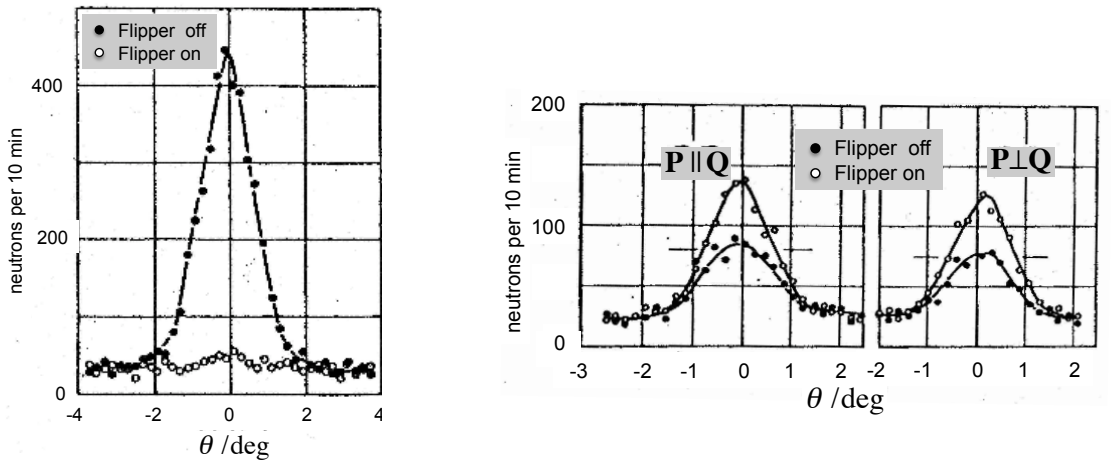


Fig. 11: Left*: Nuclear isotopic incoherent scattering from nickel obtained by rocking the analyzer crystal through the elastic position, which is essentially all non-spin-flip scattering.*
Right: *Nuclear spin-incoherent scattering from vanadium show 2/3 and 1/3 contributions in the spin-flip and non-spin-flip channel respectively. There is no dependence on the direction of* $\mathbf{P}$ *relative to* $\mathbf{Q}$ *for all nuclear scattering (from Ref.[14]).*

Applications to local order in disordered hydrogeneous materials

Typical soft matter samples contain hydrogen which causes a huge spin-incoherent background ($\sigma_{inc}(H) = 80\,\mathrm{b}$) in the wide-angle scattering that contains information about local correlations ($\sigma_{coh}(H) = 1.76\,\mathrm{b}$). Here, a precise determination of coherent scattering can be achieved by measuring spin-flip and non-spin-flip scattering. It is particularly valuable to combine this further with the method of contrast variation using H and D isotopes, having rather distinct scattering lengths, $b_{coh}(H) = -0.374 \cdot 10^{-12}\mathrm{cm}$ and $b_{coh}(D) = 0.667 \cdot 10^{-12}\mathrm{cm}$. Fig. 12 shows the separated coherent scattering of a polymer glass. Such results provide most useful information about local order that can be compared to molecular dynamics simulations of theoretical polymer models [20].

Applications to dynamics in liquids

Since in a liquid all atoms are moving around, the scattering is not elastic as in the case of Bragg peaks from a solid, single crystal. Diffraction - the energy integrated scattering - provides us with structural properties from a snap-shot of typical atomic configurations. Since neutron energies are comparable to thermal energies involved in atomic motions, it is relatively simple to achieve an adequate energy resolution to study the dynamics for instance in liquids. Therefore, a typical instrument set-up uses the time-of-flight technique: the monochromatic beam is pulsed by a mechanical chopper and the measured *time-of-flight* of the neutrons can be related to an energy transfer in the sample. Note, the separation by polarization analysis in *coherent* scattering and spin-*incoherent* scattering distinguishes *pair-correlations* from *single particle correlations*, respectively. The following example of liquid sodium [21] demonstrates in a very instructive way the complementary information that can be obtained. From simple liquid models one expects that the incoherent scattering has a Lorentzian shape in energy at constant Q, related to exponential relaxations in time, with a width that for the macroscopic limit, $Q \to 0$, is related to the macroscopic diffusion constant. On the other hand, the coherent scattering is

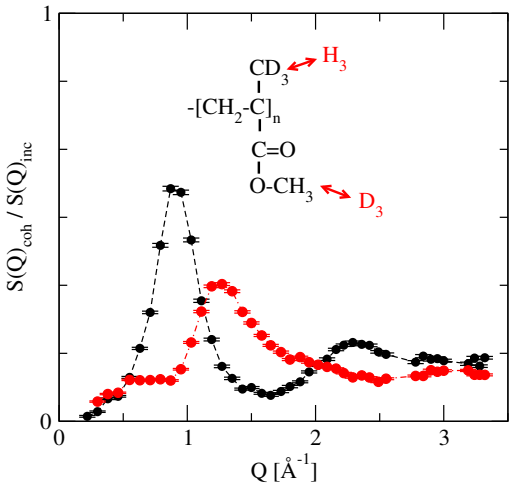


Fig. 12: *Neutron polarisation analysis separates coherent scattering from spin-incoherent scattering, which is typically a disturbing large background in materials that contain hydrogen, while here it provides a precise intrinsic calibration. In addition, H/D contrast varies the coherent scattering of a polymer glass PMMA.[20]*

rather different and exhibits a pronounced peak related to typical nearest neighbor distances reflecting precursors of Bragg peaks and crystalline order, see Fig. 13.

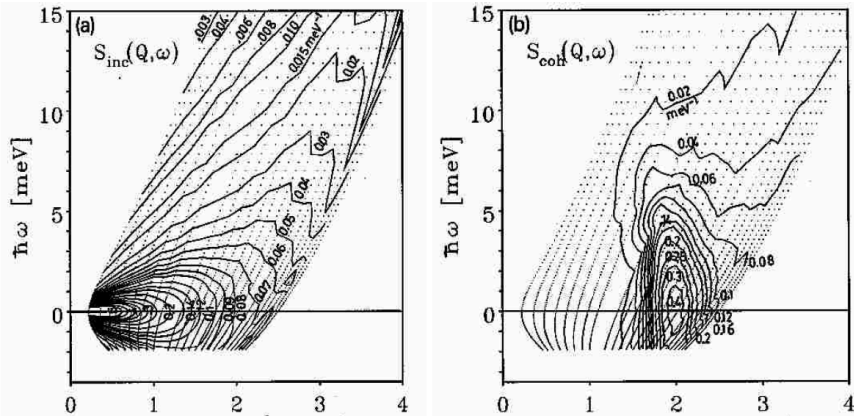


Fig. 13: *Contour plot of a) spin-incoherent and b) coherent scattering of liquid sodium at T=840 K separated by polarization analyis (taken from Ref.[21]). The dotted mesh corresponds to the coordinates of time-of-flight and scattering angles.*

The **magnetic interaction potential** is given by

$$V_m = -(\gamma_n r_0/2)\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{M}}_{\mathbf{Q}}^{\perp}, \tag{13}$$

where $\hat{\mathbf{M}}_{\mathbf{Q}}^{\perp}$ is the operator of the magnetic interaction vector,

$$\mathbf{M}_{\mathbf{Q}}^{\perp} = \mathbf{e}_{\mathbf{Q}} \times \mathbf{M}_{\mathbf{Q}} \times \mathbf{e}_{\mathbf{Q}} \tag{14}$$

which is reduced to only the perpendicular components of $\mathbf{M}_{\mathbf{Q}}$ with respect to $\mathbf{Q}$. $\mathbf{M}_{\mathbf{Q}}$ represents the total Fourier transform of the spin and orbital contribution to the magnetization density. The reason for the anisotropy of the interaction is due to the dipolar interaction of the neutron spin with the magnetic moments [19], which it is illustrated in Fig. 14. The components of a magnetic dipole field parallel to the scattering vector $\mathbf{Q}$ cancel out. Therefore, in contrast to the spin-incoherent scattering, magnetic scattering is anisotropic with respect to $\mathbf{Q}$ and only $\mathbf{M}_{\mathbf{Q}}^{\perp}$, the components perpendicular to $\mathbf{Q}$ can be observed.

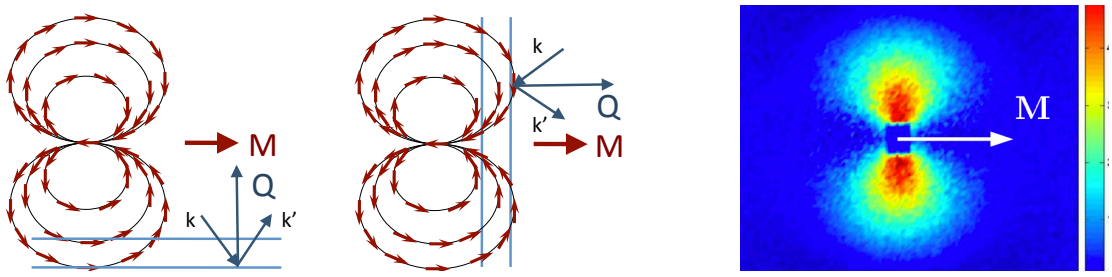


Fig. 14: *Illustration why only* $\mathbf{M}_{\mathbf{Q}}^{\perp}$ *is measured. For* $\mathbf{M} \perp \mathbf{Q}$*, magnetic dipole field amplitudes show constructive interference, for* $\mathbf{M} \parallel \mathbf{Q}$ *destructive interference.* Right: *Polarized small angle scattering probing the magnetization of iron oxide nanoparticles.[22]*

In analogy to the spin-dependent nuclear interaction, we obtain the transition matrix elements for the magnetic interaction, choosing z-polarization and x parallel to $\mathbf{Q}$, $\mathbf{M}_{x,\mathbf{Q}}^{\perp} = 0$, and

$$\langle + |\, \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{M}}_{\mathbf{Q}}^{\perp} | + \rangle = \mathbf{M}_{z,\mathbf{Q}}^{\perp}, \tag{15}$$

$$\langle - |\, \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{M}}_{\mathbf{Q}}^{\perp} | + \rangle = \mathrm{i}\mathbf{M}_{y,\mathbf{Q}}^{\perp}. \tag{16}$$

Hence, as illustrated in Fig. 15, the component of $\mathbf{P}$ parallel to $\mathbf{M}_{\mathbf{Q}}^{\perp}$ remains unchanged, while

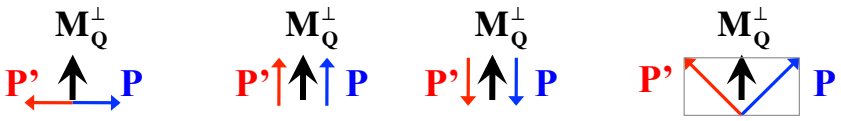


Fig. 15: *Change of initial polarization* $\mathbf{P}$ *to final polarization* $\mathbf{P}'$*: the component perpendicular to* $\mathbf{M}_{\mathbf{Q}}^{\perp}$ *reverses sign, the parallel component of* $\mathbf{P}$ *is invariant.*

the component of $\mathbf{P}$ perpendicular to $\mathbf{M}_{\mathbf{Q}}^{\perp}$ reverses its sign. This selection rule combined with the $\mathbf{Q}$-dependence provides another simple rule: If $\mathbf{P} \parallel \mathbf{Q}$, the total magnetic scattering will be spin-flip.
Therefore, as exemplified in Moon, Riste, Koehler's seminal paper [14] nuclear and magnetic Bragg peaks can be separated from non-spin-flip and spin-flip scattering respectively b scanning with $\mathbf{P} \parallel \mathbf{Q}$.

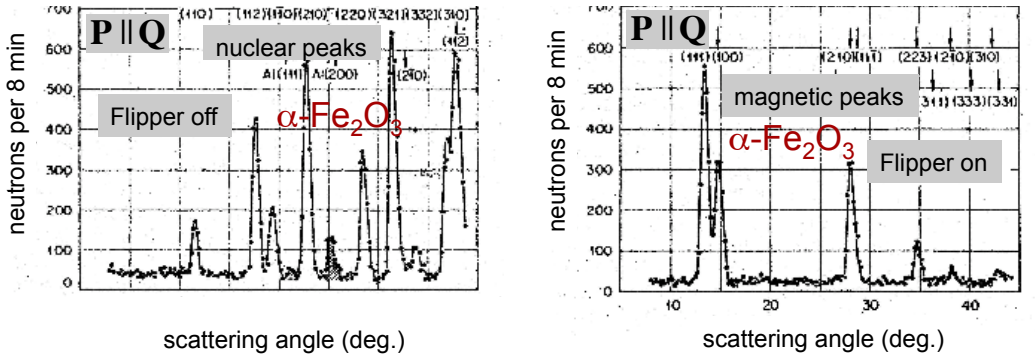


Fig. 16: *Separation of magnetic and nuclear Bragg peaks for powder diffraction from* Fe_2O_3 *by non-spin-flip and spin-flip scattering with* $\mathbf{P} \parallel \mathbf{Q}$*, from Ref.[14].*

3.2 Scattering and polarization analysis, the Blume-Maleyev equations

However, turning from the more simple expressions for the scattering amplitudes to scattering and interference of nuclear and dipolar magnetic interaction potentials, we have to face more complex expressions. As shown by Blume[23] and Maleyev[24], the scattering process can be completely described by two master equations (see Appendix). The first equation gives the scattering cross-section $\sigma_{\mathbf{Q}}$, the second one describes the final polarization $\mathbf{P}'$,

$$\begin{aligned} \sigma_{\mathbf{Q}} &= |N_{\mathbf{Q}}|^2 + |\mathbf{M}^{\perp}_{\mathbf{Q}}|^2 + \mathbf{P}(N_{-\mathbf{Q}}\mathbf{M}^{\perp}_{\mathbf{Q}} + \mathbf{M}^{\perp}_{-\mathbf{Q}}N_{\mathbf{Q}}) + \mathrm{i}\mathbf{P}(\mathbf{M}^{\perp}_{-\mathbf{Q}} \times \mathbf{M}^{\perp}_{\mathbf{Q}}) \\ \mathbf{P}'\sigma_{\mathbf{Q}} &= |N_{\mathbf{Q}}|^2\mathbf{P} + \mathbf{M}^{\perp}_{\mathbf{Q}}(\mathbf{P}\mathbf{M}^{\perp}_{-\mathbf{Q}}) + \mathbf{M}^{\perp}_{-\mathbf{Q}}(\mathbf{P}\mathbf{M}^{\perp}_{\mathbf{Q}}) - \mathbf{P}\mathbf{M}^{\perp}_{\mathbf{Q}}\mathbf{M}^{\perp}_{-\mathbf{Q}} \\ &\quad + \mathbf{M}^{\perp}_{\mathbf{Q}}N_{-\mathbf{Q}} + \mathbf{M}^{\perp}_{-\mathbf{Q}}N_{\mathbf{Q}} + \mathrm{i}\mathbf{M}^{\perp}_{\mathbf{Q}} \times \mathbf{M}^{\perp}_{-\mathbf{Q}} + \mathrm{i}(\mathbf{M}^{\perp}_{\mathbf{Q}}N_{-\mathbf{Q}} - \mathbf{M}^{\perp}_{-\mathbf{Q}}N_{\mathbf{Q}}) \times \mathbf{P} \end{aligned} \tag{17}$$

The notation uses $-\mathbf{Q}$ to denote the complex conjugate. Here, for simplicity only the $\mathbf{Q}$-dependence is specified for Bragg scattering or diffuse scattering. However, in a more general form they apply to inelastic scattering and can be related to van Hove correlation functions (see Appendix).

These equations readily show the different information that can be obtained from an unpolarized with respect to a polarized experiment. While unpolarized neutrons only measure the sum of nuclear and magnetic intensities, for polarized neutrons additional intensity may arise first, due to possible structural-magnetic (NM-terms) interference and second, due to cross products of the magnetic interaction vector $\mathrm{i}\mathbf{M}^{\perp}_{\mathbf{Q}} \times \mathbf{M}^{\perp}_{-\mathbf{Q}}$ describing chiral correlations in non-collinear spin systems. A look at the second equation for the final polarization reveals that we can identify such terms "NM" and "MxM" even with unpolarized neutrons, because they may create final polarization (set $\mathbf{P} = 0$).

3.3 Spherical neutron polarimetry

Another convenient description is given by the standard convention in spherical neutron polarimetry (SNP), or in other words spherical polarization analysis, expressing the final polarization $\mathbf{P}'$ in a tensor equation [25]

$$\mathbf{P}'\sigma \;=\; (|N|^2 + \boldsymbol{\mathcal{R}})\mathbf{P} + \mathbf{P}'' \tag{18}$$

in which the first term $(|N|^2+\boldsymbol{\mathcal{R}})\mathbf{P}$ consists of the scalar nuclear scattering $|N|^2$ and the matrix $\boldsymbol{\mathcal{R}}$ describing the rotation of $\mathbf{P}$, and $\mathbf{P}''$ is the created polarization.
Using the common specific orthogonal setting $\mathbf{x}$ parallel to $\mathbf{Q}$, and $\mathbf{y}$ and $\mathbf{z}$ perpendicular to $\mathbf{Q}$, horizontally and vertically set to the scattering plane respectively, $\boldsymbol{\mathcal{R}}$ and $\mathbf{P}$" are obtained as

$$\begin{aligned}
\boldsymbol{\mathcal{R}} &= \begin{pmatrix} -|M_y|^2 - |M_z|^2 & 2\,Im\,[NM_z] & 2\,Im\,[NM_y] \\ -2\,Im\,[NM_z] & +|M_y|^2 - |M_z|^2 & 2\,Re\,[M_yM_z] \\ -2\,Im\,[NM_y] & 2\,Re\,[M_zM_y] & -|M_y|^2 + |M_z|^2 \end{pmatrix} \\
\mathbf{P}'' &= (-2\,Im\,[M_yM_z]\,,\, 2Re[NM_y]\,,\, 2Re\,[NM_z])
\end{aligned}$$

In general, to detect all elements of the tensor $\boldsymbol{\mathcal{R}}$, one needs spherical neutron polarimetry SNP. Experimentally this can be achieved with a Cryopad device [18] using superconducting material to shield the sample area from magnetic fields. In a MuPaD device, likewise μ-metal is used obtain the zero field condition, which prevents undesired precessions. In particular, SNP is important because it allows to distinguish a rotation from a depolarization of the beam. Depolarization may occur due to an incoherent superposition of intensities with different polarization, this includes spin-incoherent scattering, and more important for determining magnetic structures, intensity from different magnetic domains. See Ref.[25] for detailed examples and analysis.
We may note that SNP does not necessarily require zero field conditions, if the precession can be controlled. Therefore, one can start with precessing polarization $\mathbf{P}$ and analyze the non-precessing polarization $\mathbf{P}'$ along the field. This provides a relatively simple principle, which in contrast to zero-field methods works even for multi-detectors and time-of-flight instruments[26, 27], see also Appendix A.2.

3.4 Uniaxial neutron polarization analysis

Most polarized neutron work is performed on TAS instruments with only uniaxial polarization analysis, sometimes also called longitudinal polarization analysis, which limits the information to the trace of the polarization transfer tensor. Typical applications are like those demonstrated in the experiments by Moon, Riste and Koehler, the separation of magnetic scattering and its directional dependence, and of nuclear coherent and nuclear spin-incoherent scattering.

Half-polarized experiments

Nuclear (or structural) -magnetic (NM) interference and chiral interference can be determined not only by SNP from the off-diagonals of $\boldsymbol{\mathcal{R}}$ but experimentally much simpler in "half-polarized" experiments, by reversing the direction of $\mathbf{P}$.

$$\begin{aligned}
\sigma_{\mathbf{Q}}(\mathbf{P}) - \sigma_{\mathbf{Q}}(-\mathbf{P}) &= 2\mathbf{P}(N_{-\mathbf{Q}}\mathbf{M}^{\perp}_{\mathbf{Q}} + \mathbf{M}^{\perp}_{-\mathbf{Q}}N_{\mathbf{Q}}) + 2\mathrm{i}\mathbf{P}(\mathbf{M}^{\perp}_{-\mathbf{Q}} \times \mathbf{M}^{\perp}_{\mathbf{Q}}) \\
&= -2\,Im\,[M_yM_z]_x\,,\, 2Re[NM_y]\,,\, 2Re\,[NM_z] \text{ for } \mathbf{P} = P_x,\; P_y, \text{and } P_z
\end{aligned}$$

The possible three terms are distinguished by the direction of $\mathbf{P}$ alone and there is no need for polarization analysis. Because of the orientation of $\mathbf{P}$ to M_y and M_z, chirality is seen in spin-flip and nuclear-magnetic interference in non-spinflip mode. NM interference can follow from either accidental coherence, external fields or inherent correlations. Applying magnetic fields in paramagnets, the NM terms measured at Bragg peaks yields a susceptibility for the atomic sites of the crystal. The example of magnetization distribution in iron oxid nanoparticles, see Fig. 14, has been determined with this approach from the nuclear magnetic interference rather than from the magnetic scattering $|\mathbf{M}_{\mathbf{Q}}^{\perp}|^2$.

XYZ-polarization analysis for isotropic samples, paramagnets and powders

Recall that by measuring with $\mathbf{P} \parallel \mathbf{Q}$ and $\mathbf{P} \perp \mathbf{Q}$, one separates and distinguishes nuclear coherent, sin-incoherent and magnetic scattering, It is straightforward to generalize the separation for magnetic powder diffraction for the use of multi-detectors, although in this case it is not possible to set the polarization parallel to all $\mathbf{Q}$ simultaneously. Therefore, one uses the method of **xyz-polarization analysis** [28], and measurements of spin-flip and non-spin-flip intensities are taken with the polarization set into three orthogonal directions, say with z perpendicular to the scattering plane. By the two horizontal polarization settings, we collect the sum of the in-plane and out-of plane magnetic intensities. The following linear combinations eliminate all nuclear scattering contributions and yield the magnetic scattering only[28]. The pre-condition is isotropy, valid for powder samples, and the method is applicable to paramagnets and antiferromagnetic ordered systems (ferromagnets cause depolarization).

$$\frac{d\sigma}{d\Omega}_{pm} = 2\left(\frac{d\sigma}{d\Omega}^{SF}_{\mathrm{x}} + \frac{d\sigma}{d\Omega}^{SF}_{\mathrm{y}} - 2\frac{d\sigma}{d\Omega}^{SF}_{\mathrm{z}}\right) = -2\left(\frac{d\sigma}{d\Omega}^{NSF}_{\mathrm{x}} + \frac{d\sigma}{d\Omega}^{NSF}_{\mathrm{y}} - 2\frac{d\sigma}{d\Omega}^{NSF}_{\mathrm{z}}\right) , \qquad (19)$$

Fig. 17 shows the XYZ-separation results for the magnetic diffuse scattering from a molecular magnet.

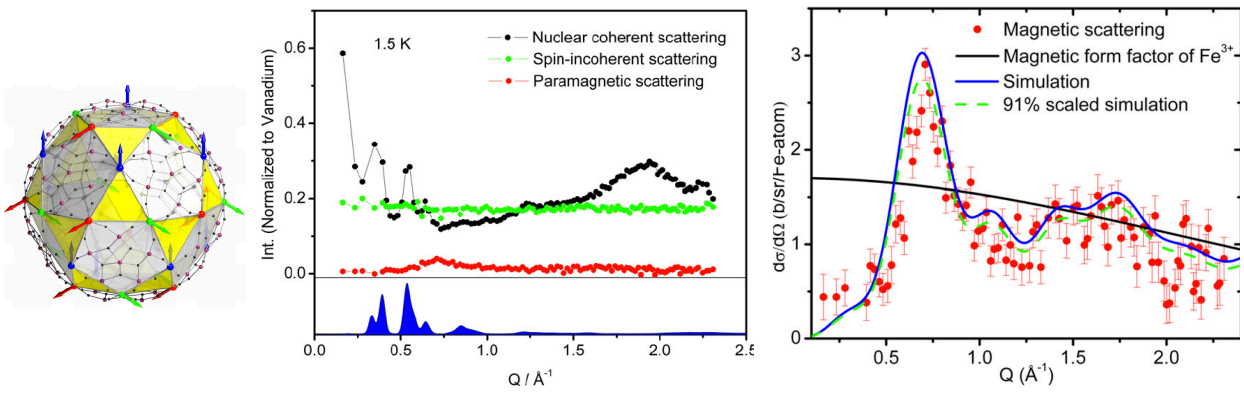


Fig. 17: *$Mo_{72}Fe_{30}X$ molecule (magnetic Fe-ions at vertices, Mo-purple; X: O-black, H, C not shown) and a 3-sublattice non-collinear spin model resulting from AF Heisenberg exchange* (left). *XYZ-separation of weak magnetic intensities* (middle). *Comparison of the magnetic intensity and the spin-model calculation* (right).*[29]*

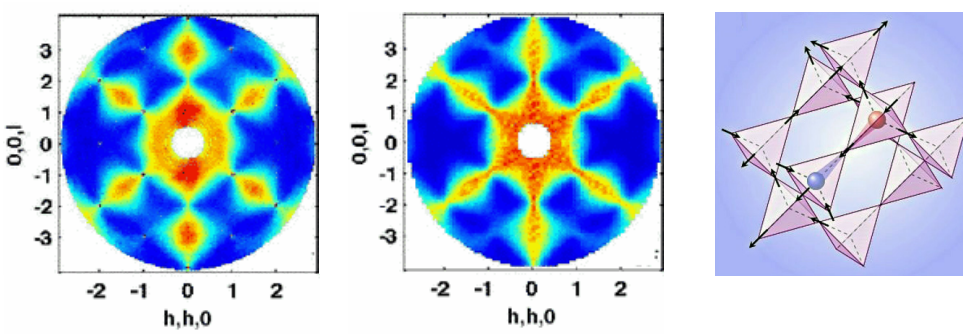


Fig. 18: *Magnetic monopoles evidenced in the diffuse scattering from spin ice, $Ho_2Ti_2O_7$, (*middle*) experimental SF scattering at T = 1.7 K with pinch points at (0,0,2), (1,1,1), and (2,2,2); (*right*) Monte Carlo simulations of the near neighbour model. [31]*

XYZ-polarization analysis for single crystals

For scattering from single crystals all terms in the Blume-Maleyev equations need to be considered. However, there is a new approach, which provides a valid XYZ-polarization analysis for single crystals, making a full separation from the diagonal terms alone by including polarization reversal, see Ref.[30]. The method also applies to inelastic scattering. The premise is the absence of different magnetic domains. Hence, a valid application is the diffuse scattering of homogeneous states in disordered phases, where the efficient use of multi detectors is particularly valuable.

The example on a "spin-ice" system $Ho_2Ti_2O_7$, a cubic pyroclore structure, Fig. 18, shows a measurement of a single polarization element, spin-flip in vertical polarization, which gives essentially the correlations of the in-plane magnetic components. The tetrahedral network and Ising $\langle 111 \rangle$ spin-anisotropy leads to strong frustration for ferro-type exchange. In the ordered state the local spin-correlations can be described for each tetrahedra by a simple rule: two spins are pointing along the $\langle 111 \rangle$ cube diagonals towards the center of the tetrahedra and two spins point outwards. Actually this rule is the perfect analogue to the ice rules in hexagonal ice, describing the hydrogen bonds around the tetrahedral environment of the O ions. Hence Pauling's famous ice model also explains why there should be a residual entropy due to remaining disorder in spin-ice, which is the origin of the broad diffuse scattering at low temperatures. The extraordinary features of this diffuse scattering are so-called pinch-points, the saddle-points in intensity at (111) and (200) positions; on one hand the intensity variation radially, along the modulus of Q, is rather smooth, involving short-range correlations, on the other hand the transverse variation at constant Q is almost discontinuous and singular, which involves many Fourier coefficients and long-range correlations. The explanation is that the ice-topology creates effectively long-range interactions, – any local decision for a specific two-in two-out spin configuration imposes far-reaching constraints for the other tetrahedra –, an effective interaction that can be mapped to Coulomb interaction between monopoles and provides a picture, where the dipole moments in their local sums over four tetrahedral sites can be viewed as two separated monopoles.[31]

A Appendices

A.1 Density matrix formulation

Properties of Pauli spin matrices

$\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, $\quad \hat{\sigma}_{\rm x} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_y = \begin{pmatrix} 0 & -{\rm i} \\ {\rm i} & 0 \end{pmatrix}, \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
$\hat{\sigma}_\alpha \hat{\sigma}_\beta = \delta_{\alpha\beta} + {\rm i} \sum_\gamma \epsilon_{\alpha\beta\gamma} \hat{\sigma}_\gamma$
$Tr(\hat{\sigma}_\alpha \hat{\sigma}_\beta) = \delta_{\alpha\beta}$
$Tr(\hat{\sigma}_\alpha \hat{\sigma}_\beta \hat{\sigma}_\gamma) = 2{\rm i}\, \epsilon_{\alpha\beta\gamma}$
$Tr(\hat{\sigma}_\alpha \hat{\sigma}_\beta \hat{\sigma}_\gamma \hat{\sigma}_\delta) = 2(\delta_{\alpha\beta}\delta_{\gamma\delta} - \delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma})$

Neutron polarization density matrix

The 2×2 neutron polarization density matrix $\hat{\rm P}$ is a linear combination of the identity matrix and the Pauli spin matrices, $\hat{\rm P} = \frac{1}{2}\left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) + \boldsymbol{P}\hat{\boldsymbol{\sigma}} \right\}$, with the polarization vector $\boldsymbol{P} = \langle \hat{\boldsymbol{\sigma}} \rangle$
In the eigenstates of $\hat{\sigma}_z$,
$\hat{\rm P} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ for spin up state $|\uparrow\rangle = |+\rangle$,
$\hat{\rm P} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for spin down state $|\downarrow\rangle = |-\rangle$,
$\hat{\rm P} = \frac{1}{2}\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ means unpolarized.
$\boldsymbol{P} = \langle \hat{\boldsymbol{\sigma}} \rangle = Tr(\hat{\rm P}\hat{\boldsymbol{\sigma}}) = \frac{1}{2}\{Tr(\hat{\boldsymbol{\sigma}}) + \boldsymbol{P} Tr(\hat{\boldsymbol{\sigma}}\hat{\boldsymbol{\sigma}})\}$.
$\langle \hat{\rm P} \rangle = Tr(\hat{\rm P}\hat{\rm P}) = \frac{1}{2}(1 + |\boldsymbol{P}|^2)$.
Boundaries for unpolarized and polarized states $\frac{1}{2} \leq |\hat{\rm P}| = |\hat{\boldsymbol{\sigma}}| \leq 1$ and $0 \leq |\boldsymbol{P}| \leq 1$.

Polarized neutron scattering cross section

Differential inelastic scattering cross section for polarized neutrons

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{k'}{k} \sum_{\lambda,\sigma} p_\lambda p_\sigma \sum_{\lambda',\sigma'} \langle \lambda, \sigma | \hat{\rm V}^+_{\mathbf{Q}} | \lambda', \sigma' \rangle \langle \lambda', \sigma' | \hat{\rm V}_{\mathbf{Q}} | \lambda, \sigma \rangle \delta(E_\lambda - E_{\lambda'} + \hbar\omega)$$

$$\hat{\rm V}_{\mathbf{Q}} = \hat{b}_{\mathbf{Q}} + \hat{\boldsymbol{\sigma}} \hat{\boldsymbol{\alpha}}_{\mathbf{Q}}, \text{with } \hat{b}_{\mathbf{Q}} = \sum_m b_m e^{-i\mathbf{Q}\cdot\mathbf{R}_m} \text{ and } \hat{\boldsymbol{\alpha}}_{\mathbf{Q}} = \sum_m B_m \hat{\boldsymbol{\sigma}} \hat{\mathbf{I}}_m e^{-i\mathbf{Q}\cdot\mathbf{R}_m} + \frac{\gamma r_0}{2\mu_B} \hat{\boldsymbol{\sigma}} \hat{\mathbf{M}}^\perp_{\mathbf{Q}}$$

$$\sum_{\sigma,\sigma'} p_\sigma \langle \sigma | \hat{\rm V}^+_{\mathbf{Q}} | \sigma' \rangle \langle \sigma' | \hat{\rm V}_{\mathbf{Q}} | \sigma \rangle = \sum_\sigma p_\sigma \langle \sigma | \hat{\rm V}^+_{\mathbf{Q}} \hat{\rm V}_{\mathbf{Q}} \hat{\rm P} | \sigma \rangle = Tr_\sigma(\hat{\rm V}^+_{\mathbf{Q}} \hat{\rm V}_{\mathbf{Q}} \hat{\rm P}) = Tr_\sigma(\hat{\rm P} \hat{\rm V}^+_{\mathbf{Q}} \hat{\rm V}_{\mathbf{Q}})$$

$$Tr_\sigma(\hat{\rm P} \hat{\rm V}^+_{\mathbf{Q}\lambda'\lambda} \hat{\rm V}_{\mathbf{Q}\lambda\lambda'}) = \frac{1}{2} Tr_\sigma \left\{ (1 + \mathbf{P}\hat{\boldsymbol{\sigma}})(\hat{b}^+_{\mathbf{Q}\lambda'\lambda} + \hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\sigma}})(\hat{b}_{\mathbf{Q}\lambda\lambda'} + \hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'}\hat{\boldsymbol{\sigma}}) \right\}$$

$$= \hat{b}^+_{\mathbf{Q}\lambda'\lambda} \hat{b}_{\mathbf{Q}\lambda\lambda'} + \hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}\lambda'\lambda} \hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'} + \hat{b}^+_{\mathbf{Q}\lambda'\lambda} (\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'} \cdot \mathbf{P}) + (\hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}\lambda'\lambda} \cdot \mathbf{P}) \hat{b}_{\mathbf{Q}\lambda\lambda'} + {\rm i}\mathbf{P} \cdot (\hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}\lambda'\lambda} \times \hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'})$$

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{k'}{k} \left([\hat{b}^+_{\mathbf{Q}} \hat{b}_{\mathbf{Q}}]_\omega + [\hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}} \cdot \hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega + [\hat{b}^+_{\mathbf{Q}} (\hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}} \cdot \mathbf{P})]_\omega + [(\hat{\boldsymbol{\alpha}}_{\mathbf{Q}} \cdot \mathbf{P}) \hat{b}_{\mathbf{Q}}]_\omega + {\rm i}\mathbf{P} \cdot [\hat{\boldsymbol{\alpha}}^+_{\mathbf{Q}} \times \hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega \right)$$

$$\left[\hat{\rm A}^+_{\mathbf{Q}} \hat{\rm B}_{\mathbf{Q}} \right]_\omega = S_{\rm AB}(\mathbf{Q}, \omega) = \frac{1}{2\pi} \int dt\, \exp({\rm i}\omega(t)) \langle \hat{\rm A}^+_{\mathbf{Q}}(t) \hat{\rm B}_{\mathbf{Q}}(0) \rangle$$

$$\text{Polarization of the scattered beam} \quad \mathbf{P'} = \langle \hat{\boldsymbol{\sigma}} \rangle = \frac{Tr(\hat{\mathrm{P}}\hat{\mathrm{V}}^{+}_{\mathbf{Q}}\hat{\boldsymbol{\sigma}}\hat{\mathrm{V}}_{\mathbf{Q}})}{Tr(\hat{\mathrm{P}}\hat{\mathrm{V}}^{+}_{\mathbf{Q}}\hat{\mathrm{V}}_{\mathbf{Q}})}$$

$$\begin{aligned}
Tr_\sigma(\hat{\mathrm{P}}\hat{\mathrm{V}}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\sigma}}\hat{\mathrm{V}}_{\mathbf{Q}\lambda\lambda'}) &= \frac{1}{2}Tr_\sigma\left\{(1+\mathbf{P}\hat{\boldsymbol{\sigma}})(\hat{b}^{+}_{\mathbf{Q}\lambda'\lambda}+\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\sigma}})\hat{\boldsymbol{\sigma}}(\hat{b}_{\mathbf{Q}\lambda\lambda'}+\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'}\hat{\boldsymbol{\sigma}})\right\} \\
\text{(nuclear)} &= \hat{b}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{b}_{\mathbf{Q}\lambda\lambda'}\mathbf{P} \\
\text{(magnetic)} &\quad + \hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}(\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'}\cdot\mathbf{P}) + (\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\cdot\mathbf{P})\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'} - \mathbf{P}(\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'}) \\
\text{(nuclear}-\text{magnetic NM)} &\quad + \hat{b}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'} + \hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{b}_{\mathbf{Q}\lambda\lambda'} \\
\text{(chiral)} &\quad - \mathrm{i}\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\times\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'} \\
\text{(rotation by NM)} &\quad + \mathrm{i}\mathbf{P}\times[\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{b}_{\mathbf{Q}\lambda\lambda'} - \hat{b}^{+}_{\mathbf{Q}\lambda'\lambda}\hat{\boldsymbol{\alpha}}_{\mathbf{Q}\lambda\lambda'}]
\end{aligned}$$

$$\begin{aligned}
\mathbf{P'}\frac{d^2\sigma}{d\Omega dE'} &= \frac{k'}{k}(\ \mathbf{P}[\hat{b}^{+}_{\mathbf{Q}}\hat{b}_{\mathbf{Q}}]_\omega + [\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}}(\mathbf{P}\cdot\hat{\boldsymbol{\alpha}}_{\mathbf{Q}})]_\omega + [\hat{\boldsymbol{\alpha}}_{\mathbf{Q}}(\mathbf{P}\cdot\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}})]_\omega - \mathbf{P}[\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}}\cdot\hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega \\
&\quad +\ [\hat{b}^{+}_{\mathbf{Q}}\hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega + [\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}}\hat{b}_{\mathbf{Q}}]^{+}_\omega + \mathrm{i}\cdot[\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}}\times\hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega \\
&\quad +\ \mathrm{i}\mathbf{P}\times[\hat{\boldsymbol{\alpha}}^{+}_{\mathbf{Q}}\hat{b}_{\mathbf{Q}} - \hat{b}^{+}_{\mathbf{Q}}\hat{\boldsymbol{\alpha}}_{\mathbf{Q}}]_\omega\ )
\end{aligned}$$

A.2 Precession spherical neutron polarimetry

Although scattering terms of all correlation functions appear on the trace of the polarization tensor, and can be separated from the twelve measurements with spinflip, non-spinflip and field reversal,[30] the off-diagonal terms are of interest in case of magnetic domains. In view of the instrumentation at the new spallation sources, one should note that multi-detector and time-of flight instruments can not apply zero field techniques for SNP, however, a precession technique for SNP has been developed for this purpose and been demonstrated at on the DNS instrument [26, 27]. Most important to note is that by measuring the non-precessing final component, one can readily measure and analyze the polarization of the scattering simultaneously for all $\mathbf{Q}$ and ω.

The principle idea is to start with precessing incident polarization and to analyze the non-precessing final polarized scattering. Therefore, (i) a $\pi/2$ flipper set the polarization into precession mode, (ii) a variation of the field strength before the sample determines the precession angle and the polarization at the sample, (iii) a variation of the field direction before the sample will rotate the precession plane to access all matrix elements; a simulation and an experiment are shown in Fig. 19 and Fig. 20.

One of the early successful examples of spherical polarization analysis solving a complex antiferromagnetic structure is the study of $U_{14}Au_{51}$ by zero-field polarimetry on CryoPad[32]. It is a magnetic structure with zero propagation vector, for which nuclear and magnetic Bragg peaks coincide. We repeated the experiment on DNS with the new precession technique. Fig. 20 illustrates the identification of nuclear and magnetic scattering for the (201) Bragg peak. The crystal is oriented vertically with its b-axis parallel to z. The spins are ordered within the a-b plane with a hexagonal star-like structure. The nuclear scattering is confirmed by a measurement above the Neel temperature at 30K and remains independent of the polarisation. Maxima are found for a rotation of the polarisation from the z to the y direction. For an initial polarisation along x

(which is set parallel to the scattering vector) only spin-flip scattering is observed for the magnetic scattering, as it must be. For all cases, when taking nuclear scattering into account, an effective depolarisation is observed. These results are in excellent agreement with the results obtained with CryoPAD [32].

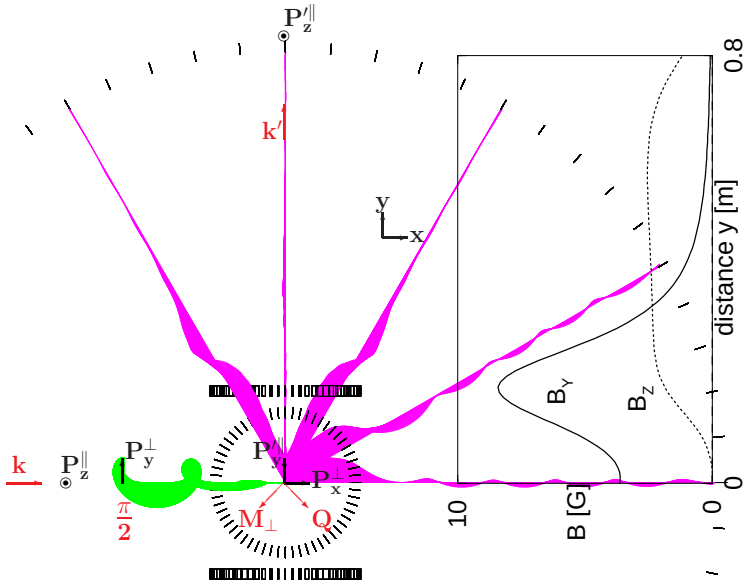


Fig. 19: *Simulation of precessing neutron polarisation analysis for typical parameters of the DNS instrument. Neutrons with wave-vector $\vec{k}$ are (i) polarized in P_z, (ii) undergo a $\pi/2$-flip to P_y, and (iii) precess to the sample to P_x in a field turning the normal vector of the precessing plane from z to y. Scattering is simulated for five different scattering angles, and in each case the final polarization turns up adiabatically from P_y to P_z. The inset shows the field variation for 90^o scattering.*

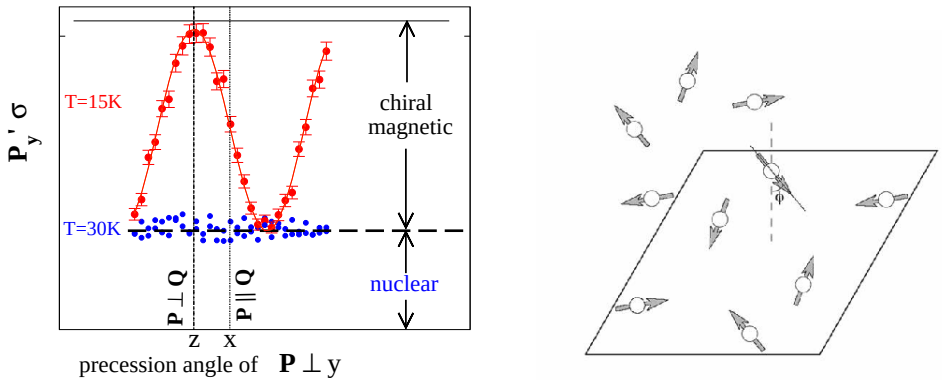


Fig. 20: *Spherical polarisation analysis on the (201) Bragg peak of $U_{14}Au_{51}$. The initial polarisation precesses perpendicular to the applied field in y-direction. At 30 K there is only a nuclear Bragg component which continues to precess after the scattering event. At lower temperatures, T=15 K, there is an additional magnetic contribution for which the neutron polarisation changes from* $\mathbf{P} \parallel z$ *to* $\mathbf{P}' \parallel y$.

For interested readers in polarized neutron studies, I like to recommend particularly the lectures of Andrey Zheludev [33], giving a more thorough discussion and a variety of examples for applications, as well as the lectures and publications of Otto Schärpf [21], which can be found on his homepage.

References

[1] Alvarez, L. W., F. A. Bloch, Quantitative determination of the neutron magnetic moment in absolute nuclear magnetons, Phys. Rev. **57**, 111 (1940).

[2] First determination of the negative sign of γ_n, P.N. Powers, The magnetic scattering of neutrons, Phys. Rev. **54**, 827 (1938).

[3] Charles M. Sommerfield, Phys. Rev. **107**, 328 (1957).

[4] Julian Schwinger, Phys. Rev. **73**, 416 (1948).

[5] G. P. Felcher, S. Adenwalla, V. O. De Haan and A. A. Vav Well, Nature **377**, 409 (1995).

[6] O. Schärpf, in *Neutron Spin Echo* Lecture Notes in Physics **128**, Ed. F. Mezei, Springer, pp. 27-52 (1980).

[7] F. Klein, *The mathematical theory of the top*, Princeton 1897; see also H. Goldstein, *Klassische Mechanik*, Akademische Verlagsges. Frankfurt am Main (1974).

[8] O. Schärpf, J. Physics E: Sci. Instrum. **8**, 268 (1975); F. Mezei, Commun. Phys. **1**, 81 (1976); P. Böni, J. Neutron Research **5**, 63 (1996).

[9] http://www.swissneutronics.ch/products/coatings.html

[10] E. Babcock *et al.*, www.jcns.info/He3_SEOP.

[11] M. T. Rekveldt, T. Keller and R. Golub, Europhys. Lett. **54**, 342 (2001).

[12] J. Repper *et al.* JCPDS-International Centre for Diffraction Data (2009) ISSN 1097-0002.

[13] P.G. Niklowitz *et al.*, Phys. Rev. Lett. **104**, 106406 (2010).

[14] R.M. Moon, T. Riste, W.C. Koehler, Phys. Rev. **181**, 920 (1969).

[15] W. Schweika and P. Böni, Physica B **297**, 155 (2001).

[16] J. R. Stewart *et al*, J. Appl. Cryst. **42**, 69 (2009).

[17] O. Schärpf, Physica B **182**, 376 (1992).

[18] F. Tasset *et al.*, Physica B **267** (1-4) 69-74 (1999).

[19] G.L. Squires, *Introduction to the theory of thermal neutron scattering*, Cambridge University Press, Cambridge (1978).

[20] A. C. Genix *et. al.* Macromolecules **39**, 3947 (2006).

[21] O. Schärpf, *The spin of the neutron as a measuring probe*, Chapter 11, http://82.135.31.182/neutronpol.pdf

[22] S. Disch *et al.*, New Journal of Physics **14**, 013025 (2012).

[23] M. Blume, Phys. Rev. **130**, 1670 (1963).

[24] S. V. Maleyev, V. G. Baryakhtar P. A. and Suris, Fiz. Tv. Tela **4**, 3461 (1962); Soviet Phys. Solid State **4**, 2533 (1963).

[25] P. J. Brown, J. B. Forsyth, and F. Tasset, Proc. Roy. Soc. London A **442**, 147 (1993); P.J. Brown, *Polarimatric Neutron Scattering*, in lecture notes on *Polarized Neutron Scattering* (Eds. Th. Brückel and W. Schweika), Schriften des Forschungszentrums Jülich, Series Matter and Materials Vol.**12**, Forschungszentrums Jülich (2002).

[26] W. Schweika, Physica B **335** (1-4), 157 (2003).

[27] W. Schweika, S. Easton, and K.-U. Neumann, Neutron News **16**, 14 (2005).

[28] O. Schärpf, H. Capellmann, Phys. Stat. Sol. (a) **135**, 359 (1993).

[29] Z. Fu *al.*, New Journal of Physics **12**, 083044 (2010).

[30] W. Schweika, J. Phys. Conf. Ser. **211**, 012026 (2010).

[31] T. Fennell *et al.* Science **326**, 415 (2009).

[32] P. J. Brown, J. Crangle, K.-U. Neumann, J. G. Smith, and K. R. A. Ziebeck, J. Phys.: Cond. Mat. **9**, 4729 (1997).

[33] www.neutron.ethz.ch/education/Lectures/neutronspring.

C 7 Polarization Handling: Synchrotron [1]

Shibabrata Nandi
Institut für Festkörperforschung
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Investigations of the structure and dynamics of materials have been an important and essential endeavor in condensed matter physics since the early 20^{th} century. Both neutron and x-ray scattering techniques have been used extensively to study the crystallographic structure of materials and provide complementary views of structure. For example, neutron scattering has traditionally been the standard tool for studies of magnetic structure and the dynamics of condensed matter systems. X-ray diffraction has largely been applied to detailed crystallographic structure determination. The principle interaction that makes structure determination possible for x-rays is the Coulomb interaction between x-rays and the electronic distribution which gives rise to driven harmonic oscillation of the electrons, and the emission of electric dipole radiation. This is the classical *Thomson scattering* process.

At x-ray absorption edges photoelectric absorption occurs and electrons are promoted from core levels into empty states above the Fermi level. Photons that take part in the photoelectric absorption are lost for the scattering experiment. However, the incident photons can also give rise to virtual transitions between core levels and states above the Fermi level that relax back to the core states with the emission of x-rays with the same energy as the initial beam. For charge scattering, this is known as *anomalous charge scattering* which yields additional terms in the x-ray scattering form factor that can be used to enhance the scattering contrast between neighboring elements. Anomalous scattering is also sensitive to the anisotropy of local environment, such as the arrangement of orbitals and orbital order. In addition to charge scattering, x-rays interact with the magnetic moment of the system. Indeed, the *magnetic scattering* of x-rays from electrons in molecules and solids is well documented in theory [1] and was observed experimentally by de Bergevin and Brunel in 1970's with a commercial x-ray tube [2]. de Bergevin and Brunel also presented a classical picture of the interaction between x-rays with electrons and magnetic moments [3] illustrated in Fig. 1. Unfortunately, from a practical point of view, the x-ray scattering cross-sections from electron spins are reduced by approximately six orders of magnitude compared to normal charge scattering and, therefore, using x-ray magnetic scattering to study magnetism was largely impractical in 1970's.

The situation changed in 1990's when Gibbs and coworkers successfully observed magnetic scattering from Ho metal due to the dramatic increase of photon flux available from synchrotron radiation.[4] The increased photon flux compensates for the weakness of the magnetic scattering signal. The polarization properties and tunability of the x-ray energy at a synchrotron source provides additional advantages for magnetic x-ray scattering. As will be discussed in the next section when x-ray energies are tuned through the absorption edges of an element of interest there is a resonant enhancement of the scattering signal now known as *x-ray resonant magnetic scattering* [5]. Away from the resonance condition, the magnetic scattering signal is known as *non-resonant x-ray magnetic scattering*. All of the above processes have been reviewed and described in detail in several texts [6, 7]. In the following, we will concentrate mainly on the basic principles and applications of the x-ray resonant magnetic scattering and nonresonant magnetic scattering.

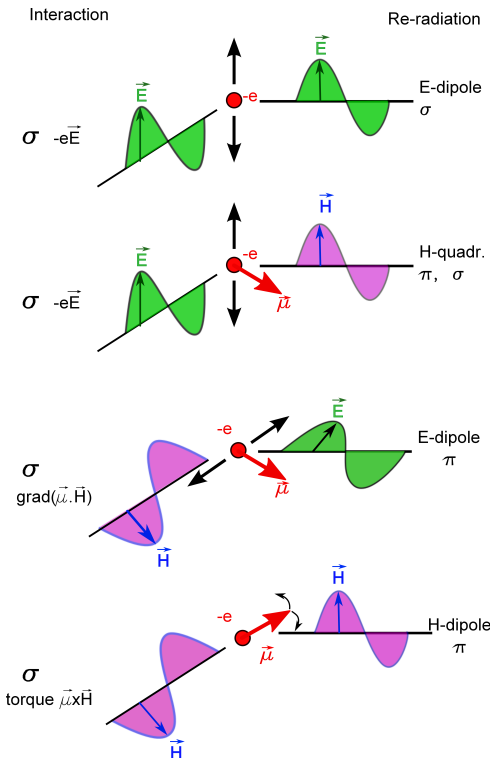


Figure 1: *Illustration of the processes leading to scattering of x-rays by the charge (top) and the spin moment (bottom three) of the electron in a classical picture. Figure has been adapted from de Bergevin and Brunel [3] and Th. Brückel [8].*

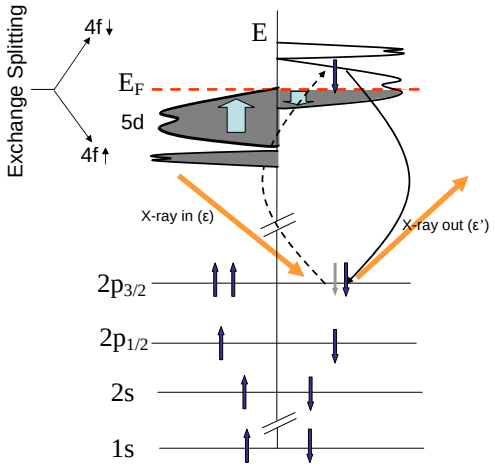


Figure 2: *Schematic illustration of the second order perturbation process leading to XRMS in the case of a lanthanide metal. An electron being photo-excited from the core level to the empty states above the Fermi energy E_F . The subsequent decay of the electron to the core level gives rise to an elastically scattered photon.*

2 Cross-section for the resonant and nonresonant magnetic x-ray scattering:

The calculation of the x-ray scattering cross-section from a quasi-relativistic Hamiltonian for a electron in a quantized electromagnetic field using second order perturbation theory was done by Blume [9] and Blume and Gibbs [10], and was later presented in a simplified form by Hill and McMorrow [11] in the coordinate system convenient for an x-ray resonant magnetic scattering experiment (XRMS) experiments. We start with the Hamiltonian for electrons in a quantized electromagnetic field:

$$\begin{aligned} H =& \sum_j \frac{1}{2m}\{\vec{P_j} - \frac{e}{c}\vec{A}(\vec{r}_{ij})\}^2 + \sum_{ij} V(\vec{r}_{ij}) - \frac{e\hbar}{2mc}\sum_j \vec{s}_j \cdot \nabla_j \times \vec{A}(\vec{r}_j) \\ & - \frac{e\hbar}{2(mc)^2}\sum_j \vec{s}_j \cdot \vec{E}(\vec{r}_j) \times \{\vec{p}_j - \frac{e}{c}\vec{A}(\vec{r}_j)\} + \sum_{\vec{k}\lambda} \hbar\omega_{\vec{k}}\{c^\dagger(\vec{k}\lambda)c(\vec{k}\lambda) + \frac{1}{2}\} \end{aligned} \tag{1}$$

Here, the first term corresponds to the kinetic energy of the electrons in the electromagnetic field, represented by the vector potential $\vec{\mathrm{A}}(\vec{\mathrm{r}})$. The second term corresponds to the Coulomb interaction between electrons, the third term corresponds to the Zeeman energy, $-\vec{\mu}\cdot\vec{H}$, of the electrons with spin $\vec{\mathrm{s}}_j$, the fourth term is the spin-orbit coupling in the initial or final states and the last term is the self energy of the electromagnetic field.
The vector potential $\vec{\mathrm{A}}(\vec{\mathrm{r}})$ is linear in photon creation and annihilation operator $c^\dagger(\vec{k}\lambda)$ and $c(\vec{k}\lambda)$, and can be expanded in plane wave form as:

$$\vec{A}(\vec{r}) = \sum_{\vec{q}\sigma}(\frac{2\pi\hbar c^2}{V\omega_q})^{\frac{1}{2}} \times [\vec{\epsilon}(\vec{q}\sigma)c(\vec{q}\sigma)e^{i\vec{q}\cdot\vec{r}} + \vec{\epsilon}^\star(\vec{q}\sigma)c^\dagger(\vec{q}\sigma)e^{-i\vec{q}\cdot\vec{r}}] \tag{2}$$

Here V is the quantization volume and $\epsilon(\vec{q}\sigma)$ is the unit polarization vector corresponding to a wave vector $\vec{q}$ and polarization state σ. The index $\sigma(=1,2)$ labels two polarizations for each wave vector $\vec{q}$. Since $\vec{\mathrm{A}}(\vec{\mathrm{r}})$ is linear in photon creation and annihilation operator, $c^\dagger(\vec{k}\lambda)$ and $c(\vec{k}\lambda)$, scattering occurs in second order for terms linear in $\vec{\mathrm{A}}$ and in first order for quadratic terms. For the spin orbit term in Eq. 1, $\vec{E}$ can be written in terms of scalar potential ϕ and the vector potential $\vec{A}$ as:

$$\vec{E} = -\nabla\phi - \frac{1}{c}\dot{\vec{A}} \tag{3}$$

After substituting $\vec{A}$ and $\vec{E}$ in Eq. 1, the Hamiltonian in Eq. 1 can be re-written as the sum of three terms:[9]

$$H = H_0 + H_R + H' \tag{4}$$

Where H_0 contains only the degrees of freedom for the electron system, H_R is the Hamiltonian for the quantized electromagnetic field and H' corresponds to the interaction between the electron and the radiation field. Scattering cross sections are calculated by assuming the solid is in a quantum state $|a\rangle$ which is an eigenstate of H_0 with energy E_0 and there is a single photon present. We then calculate the probability of a transition induced by the interaction term H' to a state $|b\rangle$, with photon $\vec{k}'\lambda'$. The transition probability (Ω) per unit time can be calculated using

"Fermi's golden rule" up to second order:

$$\Omega = \frac{2\pi}{\hbar}|\langle f|H^{'}|i\rangle + \sum_{n}\frac{\langle f|H^{'}|n\rangle\langle n|H^{'}|i\rangle}{E_i - E_f}|^2 \times \delta(E_i - E_f)$$
$$|i\rangle = |a;\vec{k}\lambda\rangle;\ |f\rangle = |b;\vec{k^{'}}\lambda^{'}\rangle \tag{5}$$

2.1 X-ray resonant magnetic scattering:

The cross-section for the elastic scattering can be written as:

$$\frac{d\sigma}{d\omega} = r_0^2\left|\sum_n e^{i\vec{Q}\cdot\vec{r_n}}f_n(\overrightarrow{k},\overrightarrow{k}',\hbar\omega)\right|$$

where, $d\sigma$ is the differential cross-section of scattering into the solid angle $d\omega$, $r_0 = \frac{e^2}{mc^2} = 2.8\times10^{-15}\,m$ is the classical electron radius, $\overrightarrow{Q} = \overrightarrow{k} - \overrightarrow{k}'$ is the momentum transfer, $\overrightarrow{k}$ and $\overrightarrow{k}'$ are the incident and scattered wave-vectors of the photon, f_n is the scattering amplitude of the nthatom which is at the position $\vec{r_n}$. While the detailed derivation of the scattering cross section is given in Ref. [9], we will outline here the final results. For coherent elastic scattering, ($|a\rangle = |b\rangle$), and the amplitude can be written as a sum over the following terms:[11]

$$f = f_0 + f^{'} + if^{''} + f_{spin} \tag{6}$$

Here, $f_0 \propto Zr_0$ is the Thompson charge scattering amplitude and f_{spin} is the non-resonant spin-dependent magnetic scattering amplitude. Far from resonance, $f^{'}$ and $f^{''}$ contribute terms proportional to the orbital and spin angular momentum. At resonance both electric and magnetic multipole transitions contribute through the terms $f^{'}$ and $f^{''}$. However, the electric dipole and quadrupole transitions are dominant with respect to magnetic multipole transitions by a factor of $\hbar\omega/mc^2$ ($\sim$ 60 for typical x-ray edges) and so the only electric multipole transitions will be considered here.
For the electric 2^L-pole resonance in a magnetic ion, the resonant contribution to the coherent scattering amplitude can be written as[12]

$$f^e_{EL}(\omega) = \frac{4\pi}{k}f_D\sum_{M=-L}^{L}[\hat{\epsilon}^{'\star}\cdot Y^{(e)}_{LM}(\hat{k}^{'})Y^{(e)\star}_{LM}(\hat{k})\cdot\hat{\epsilon}]F^e_{EL}(\omega) \tag{7}$$

Where $\hat{\epsilon}$ and $\hat{\epsilon}^{'\star}$ are the incident and scattered polarization vectors, and $\hat{k}$ and $\hat{k}^{'}$ are the incident and scattered wave vectors, respectively. $Y^{(e)}_{LM}(\hat{k})$ are the vector spherical harmonics and f_D is the Debye-Waller factor. The strength of the resonance is determined by the factor $F^e_{EL}(\omega)$, which is, to 0$^{\text{th}}$ order, determined by atomic properties:

$$F^e_{EL}(\omega) = \sum_{a,n}\frac{P_aP_a(n)\Gamma_x(aMn;EL)}{2(E_n - E_a - \hbar\omega - i\Gamma/2)} \tag{8}$$

Here, $|n\rangle$ is the excited state of the ion and $|a\rangle$ the initial state. P_a is the probability of the ion existing in the initial state $|a\rangle$ and $P_a(n)$ is the probability for a transition from $|a\rangle$ to an excited state $|n\rangle$. It is determined by overlap integrals between the two states $|a\rangle$ and $|n\rangle$. Γ_x and Γ are the partial line widths of the excited state due to a pure electric 2^L-pole (EL) radiative decay and

due to all processes, both radiative and non-radiative (including, for example, Auger decay), respectively. These electric multipole (predominantly dipole and quadrupole) transitions involve the virtual photo-excitation of an electron from a core level to the unoccupied states above the Fermi energy with a subsequent de-excitation to the core-levels yielding an elastically scattered photon. These processes become sensitive to the magnetic state in exchange split bands due to the difference in occupation of minority and majority bands as illustrated schematically in Fig. 2. Due to the resonant denominator in Eq. 8, enhancements occur at the absorption edges of the magnetic elements (e.g. when, E_n-$E_a = \hbar\omega$). The strengths of these enhancements for XRMS depend mainly on three factors as discussed by Hannon *et al.* [12] and shown by XRMS experiments on a series of rare-earth intermetallic compounds ($R\mathrm{Ni_2Ge_2}$, R= rare-earths) by Kim *et al.* [13]:

- The magnitude of the transition matrix element. Dipole transitions between states $|a\rangle$ and $|n\rangle$ differing in orbital angular momentum quantum number by ΔL = 1 are generally stronger than quadrupolar transitions with ΔL = 2. A large overlap of the wave functions $|a\rangle$ and $|n\rangle$ favors large transition matrix elements. In contrast, transitions from "s" core levels to "p" or "d" excited states do not show large resonance enhancements due to the small overlap of the wave functions.

- The difference in the density of empty states above the Fermi level for minority and majority spin states. To give an example: in lanthanide metals, the $5d$ bands are spin polarized due to the magnetic $4f$ states. However, the exchange splitting in the $5d$ is much weaker as compared to the $4f$ states and dipolar transitions $2p \rightarrow 5d$ are sometimes not much stronger than quadrupolar transitions $2p \rightarrow 5f$.

- The strength of the spin-orbit coupling in the ground and excited states. It is this coupling that provides electric multipole transitions with sensitivity to the spin magnetism.

Using the above-mentioned factors, we can qualitatively categorize the possible transitions according to the magnitude of the resonance enhancement, as listed in Table. 1.[8] Here we define the term "resonant enhancement" as the ratio between the intensity of magnetic Bragg peaks at the maximum of the resonance relative to the intensity for non-resonant magnetic scattering.
One of the strengths of resonant magnetic scattering is that the polarization state of the scattered x-rays can be modified with respect to that of the incident beam. Therefore, by analyzing the polarization of scattered x-rays, it is possible to discriminate between charge and magnetic scattering. Furthermore, by analyzing the intensity of scattered x-rays in different polarization channels, the spatial components of the ordered magnetic moment can be obtained.[14, 15] Therefore, in the following sections explicit relationship between the amplitude of scattered x-rays and incident x-rays will be shown according to Ref. [11].
For rare-earth L-edges, electric dipole transitions usually dominate the resonant magnetic cross section and are the simplest to calculate. An example of such a transition is the $2p_{3/2} \rightarrow 5d_{1/2}$ transition of Ho, which occurs at the L_{III} absorption edge. At this transition, the vector spherical harmonics can be written, for L = 1, M = $\pm$1:

$$[\hat{\epsilon}'^{\star} \cdot Y^{(e)}_{1\pm1}(\hat{k}')Y^{(e)\star}_{1\pm1}(\hat{k}) \cdot \hat{\epsilon}] = (3/16\pi)[\hat{\epsilon}' \cdot \hat{\epsilon} \mp i(\hat{\epsilon}' \times \hat{\epsilon}) \cdot \hat{z}_n - (\hat{\epsilon}' \cdot \hat{z}_n)(\hat{\epsilon} \cdot \hat{z}_n)] \tag{9}$$

Similarly, for L = 1, M = 0

$$[\hat{\epsilon}'^{\star} \cdot Y^{(e)}_{10}(\hat{k}')Y^{(e)\star}_{10}(\hat{k}) \cdot \hat{\epsilon}] = (3/8\pi)[(\hat{\epsilon}' \cdot \hat{z}_n)(\hat{\epsilon} \cdot \hat{z}_n)] \tag{10}$$

Table 1: *Magnitude of the resonance enhancement for XRMS for some elements relevant for magnetism. Only order of magnitude estimates are given with "weak" corresponding to a factor of about "10^0", "medium" to about "10^2" and "strong" to "$>10^3$" compared to the non-resonant magnetic scattering. After Ref. [66]*

Elements	Edge	Transition	Energy Range (keV)	Resonance Strength	Comment
$3d$	K	$1s \rightarrow 4p$	5-9	Weak	Small overlap
$3d$	L_I	$2s \rightarrow 3d$	0.5-1.2	Weak	Small overlap
$3d$	L_{II}, L_{III}	$2p \rightarrow 3d$	0.4-1.0	Strong	Dipolar, Large overlap, high spin polarization of $3d$
$4f$	K	$1s \rightarrow 4p$	40-63	Weak	Small Overlap
$4f$	L_I	$2s \rightarrow 5d$	6.5-11.0	Weak	Small overlap
$4f$	L_{II}, L_{III}	$2p \rightarrow 5d$, $2p \rightarrow 4f$	6.0-10.0	Medium	Dipolar and quadrupolar
$4f$	M_I	$3s \rightarrow 5p$	1.4-2.5	Weak	Small overlap
$4f$	M_{II}, M_{III}	$3p \rightarrow 5d$, $3p \rightarrow 4f$	1.3-2.2	Medium to strong	Dipolar, quadrupolar
$4f$	M_{IV}, M_V	$3d \rightarrow 4f$	0.9-1.6	Strong	Dipolar, large overlap, high spin polarization of $4f$
$5f$	M_{IV}, M_{II}	$3d \rightarrow 5f$	3.3-3.9	Strong	Dipolar, large overlap, high spin polarization of $5f$

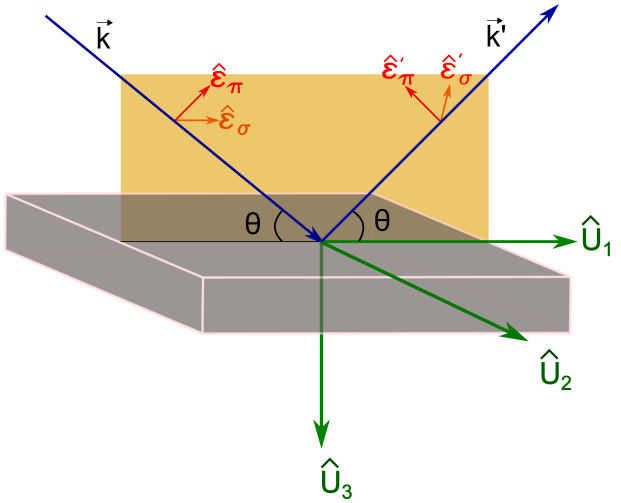


Figure 3: *The coordinate system used for the polarization dependence of the resonant scattering amplitudes described in the text. $\hat{k}$ and $\hat{k}'$ are the incident and scattered wave vectors and θ is the Bragg angle. $\hat{\epsilon}_\sigma$ ($\hat{\epsilon}'_\sigma$) and $\hat{\epsilon}_\pi$ ($\hat{\epsilon}'_\pi$) are the components of the polarization perpendicular and parallel to the scattering plane for incident (scattered) x-rays.*

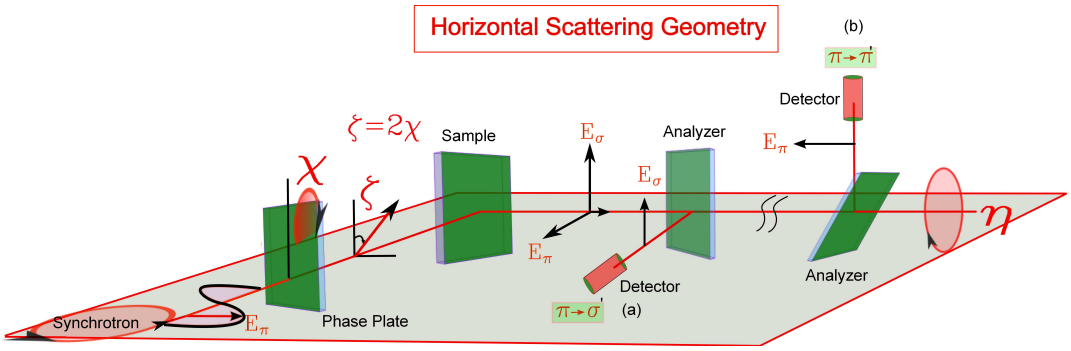


Figure 4: *Experimental realization of different polarization channels for the (a) $\pi \rightarrow \sigma'$ and (b) $\pi \rightarrow \pi'$ scattering geometries.*

where $\hat{z}_n$ is a unit vector in the direction of the magnetic moment of the nth ion. Thus,

$$f_{nE1}^{XRMS} = [(\hat{\epsilon}^{'} \cdot \hat{\epsilon})F^{(0)} - i(\hat{\epsilon}^{'} \cdot \times \hat{\epsilon}) \cdot F^{(1)} + (\hat{\epsilon}^{'} \cdot \hat{z}_n)(\hat{\epsilon} \cdot \hat{z}_n)F^{(2)}] \tag{11}$$

where $F^{(i)}$'s are the terms containing dipole matrix elements which have been evaluated by Hamrick for several rare-earths.[16] The first term of Eq. 11 contributes to the charge Bragg peak as it does not contain any dependence on the magnetic moment. The other two terms depend on the magnetic moment. All terms in Eq. 11 can be represented in 2×2 matrix form with polarization states chosen either parallel and perpendicular to the scattering plane and resolving each of the vectors $\hat{k}$, $\hat{k}^{'}$ and $\hat{z}_n$ into their components along the three orthogonal axes defined with respect to the diffraction plane shown in Fig. 3.

$$\begin{aligned} f_{nE1}^{XRMS} &= \begin{pmatrix} A_{\sigma\rightarrow\sigma'} & A_{\pi\rightarrow\sigma'} \\ A_{\sigma\rightarrow\pi'} & A_{\pi\rightarrow\pi'} \end{pmatrix} \\ &= F^{(0)} \begin{pmatrix} 1 & 0 \\ 0 & \cos 2\theta \end{pmatrix} - iF^{(1)} \begin{pmatrix} 0 & z_1 \cos\theta + z_3 \sin\theta \\ z_3 \sin\theta - z_1 \cos\theta & -z_2 \sin 2\theta \end{pmatrix} \\ &\quad + F^{(2)} \begin{pmatrix} z_2^2 & -z_2(z_1 \sin\theta - z_3 \cos\theta) \\ z_2(z_1 \sin\theta + z_3 \cos\theta) & -\cos^2\theta(z_1^2 \tan^2\theta + z_3^2) \end{pmatrix} \end{aligned}$$

Where $A_{\sigma\rightarrow\sigma'}$ represents scattering amplitude from the incident σ polarized x-rays to the scattered σ' polarized x-rays and so on.
The above matrix equation is sum of three sub-matrices: The first one contributes to the charge Bragg peak. The second and third contribute to the magnetic Bragg peak. A few points to note for the magnetic contribution:

1. The second term is linear in the components of the magnetic moment and therefore, mainly responsible for producing peaks at the magnetic wavevector, $\vec{\tau}$. For example, in an antiferromagnet with wavevector, $\vec{\tau}$, the linear term produces first harmonic satellite peaks. The third term, which is quadratic in the components of a magnetic moment produces second harmonic satellites as well as contributes to the charge Bragg peaks.[2] For a ferromagnetic material both the second and third term contribute to the positions of allowed charge reflections.

[2]for a commensurate antiferromagnet with propagation vector $\vec{\tau}$=0, it can produce intensity at the charge forbidden reciprocal lattice points.

2. Synchrotron radiation is linearly polarized in the plane of storage ring. Therefore, all combinations of polarization channels are accessible by suitably selecting the scattering plane (see Fig. 4 for experimental geometries) and, hence, different components of the ordered magnetic moment can be probed. For example, scattering in the $\sigma \rightarrow \pi^{'}$ channel is sensitive to the component of magnetic moment in the scattering plane (in the Fig. 3, z_1 and z_3) and $\pi \rightarrow \pi^{'}$ channel is sensitive to the components of a magnetic moment out of the scattering plane (z_2 in Fig. 3).

3. The scattering amplitude depends on the Bragg angle, and therefore by analyzing the magnetic peak intensities as a function of scattering angle, a specific magnetic model can be proved or disproved.

The scattering cross-section for the quadrupole transitions is much more complex and the reader is referred to the Ref. [17] for details.

2.2 Nonresonant x-ray magnetic scattering:

In many cases, the resonant process of XRMS is neither well understood nor efficient, especially at *K* edges for 3*d* transition metals. Nonresonant x-ray magnetic scattering (NRXMS) may be appropriate for these situations. The scattering amplitude of nonresonant x-ray magnetic scattering is directly related to the magnetic moment of the ions. While neutron scattering does not distinguish spin ($\vec{S}$) and orbital angular momentum ($\vec{L}$), $\vec{S}$ and $\vec{L}$ contribute differently to the scattering amplitude of the NRXMS. Far away from absorption edges of the all elements the scattering amplitude can be written as:[10]

$$f_n(\overrightarrow{Q}) = f_n^{charge}(\overrightarrow{Q}) + f_n^{NRXMS}(\overrightarrow{Q}, \overrightarrow{k}, \overrightarrow{k}')$$

In the limit of high photon energies, the nonresonant scattering amplitude may be written as

$$\begin{aligned} f_n^{NRXMS}(\overrightarrow{Q}) &= -i\frac{\hbar\omega}{mc^2}[\frac{1}{2}\overrightarrow{L_n}(\overrightarrow{Q}) \cdot \overrightarrow{A} + \overrightarrow{S_n}(\overrightarrow{Q}) \cdot \overrightarrow{B}] \\ &= -i\frac{\hbar\omega}{mc^2}\langle M_m \rangle \end{aligned} \tag{12}$$

Here $\overrightarrow{L_n}(\overrightarrow{Q})$ and $\overrightarrow{S_n}(\overrightarrow{Q})$ are the Fourier transforms of the orbital and spin magnetization densities, respectively. The vectors $\overrightarrow{A}$ and $\overrightarrow{B}$ contain different polarization and $\overrightarrow{Q}$ dependencies. Here the $\langle\rangle$ sign represent ground state expectation value of the operators. Due to this different $\overrightarrow{Q}$ and polarization dependencies in the nonresonant scattering amplitude, the orbital and spin components can be distinguished. This is in stark contrast to magnetic neutron scattering, where the scattering amplitude is sensitive to the sum of orbital and spin angular momentum. More explicitly, the expression for the neutron magnetic scattering can be written as:[9]

$$\langle M_n \rangle = \overrightarrow{Q} \times \{[\frac{1}{2}\overrightarrow{L_n}(\overrightarrow{Q}) + \overrightarrow{S_n}(\overrightarrow{Q})] \times \overrightarrow{Q}\} \cdot \overrightarrow{\sigma}$$

where $\overrightarrow{\sigma}$ is the neutron spin operator. For the point of view of performing synchrotron experiments, it is convenient to express $\langle M_m \rangle$ in a basis whose components are parallel and perpendicular to the scattering plane.

$$\begin{aligned}\langle M_m \rangle &= \begin{pmatrix} A_{\sigma\to\sigma'} & A_{\sigma\to\pi'} \\ A_{\pi\to\sigma'} & A_{\pi\to\pi'} \end{pmatrix} \\ &= \begin{pmatrix} \overrightarrow{S}\cdot(\hat{k}\times\hat{k}') & -\frac{Q^2}{2k^2}[\{\frac{\vec{L}(\vec{Q})}{2}+\overrightarrow{S}(\overrightarrow{Q})\}\cdot\hat{k}'+\frac{\vec{L}(\vec{Q})}{2}\cdot\hat{k}] \\ \frac{Q^2}{2k^2}[\{\frac{\vec{L}(\vec{Q})}{2}+\overrightarrow{S}(\overrightarrow{Q})\}\cdot\hat{k}'+\frac{\vec{L}(\vec{Q})}{2}\cdot\hat{k}'] & [\frac{Q^2}{2k^2}\overrightarrow{L}(\overrightarrow{Q})+\overrightarrow{S}(\overrightarrow{Q})]\cdot(\hat{k}\times\hat{k}') \end{pmatrix} \\ & \begin{pmatrix} (\sin 2\theta)S_2 & -2(\sin^2\theta)[(\cos\theta)(L_1+S_1)-(\sin\theta)S_3] \\ 2(\sin^2\theta)[(\cos\theta)(L_1+S_1)+(\sin\theta)S_3] & (\sin 2\theta)[2(\sin^2\theta)L_2+S_2] \end{pmatrix}\end{aligned}$$

Here we note that the diagonal elements are sensitive to the magnetization perpendicular to the scattering plane whereas the off-diagonal elements are sensitive to the magnetization within the scattering plane. Further, $\sigma \to \sigma'$ component is independent of orbital magnetization densities. In the same basis, the matrix describing the charge scattering can be written as:

$$\langle M_n \rangle = \rho(\overrightarrow{Q}) \begin{pmatrix} 1 & 0 \\ 0 & \cos 2\theta \end{pmatrix}$$

where $\rho(\overrightarrow{Q})$ is the Fourier transform of the electronic charge density.
Following Blume *et al*. [9] an estimate of the pure non resonant magnetic scattering strength compared to the charge scattering strength can be obtained as follows:

$$\frac{\sigma_{mag}}{\sigma_{charge}} = \left(\frac{\hbar\omega}{mc^2}\right)^2 \left(\frac{N_m}{N}\right)^2 \left(\frac{f_m}{f}\right)^2 \langle\mu^2\rangle$$

where $\frac{N_m}{N}$ is the ratio of the number of magnetic electrons compared to the total number of electrons, $\frac{f_m}{f}$ is the ratio between magnetic and charge form factors and $\langle\mu^2\rangle$ is the average value square of the magnetic moment. For Fe and 10 keV photons,

$$\frac{\sigma_{mag}}{\sigma_{charge}} \sim 4\times10^{-6}\langle\mu^2\rangle$$

Despite the smallness of the NRXMS compared to the charge scattering, NRXMS has been observed even for very small magnetic moment systems due to the intrinsic brilliance of the synchrotron radiation sources. [18, 19, 20, 21]

2.3 Examples of the resonance and the noresonance:

Before proceeding further, it is better to give real examples of resonant and nonresonant magnetic scattering. Figure 5 shows energy scans at the Sm L_2, L_3 and Fe K absorption edges for the compound SmFeAsO. Energy scans were performed at the magnetic Bragg positions of Sm (Fig. 5a & 5b) and Fe (Fig. 5c) moments, respectively. At 6 K (below the magnetic ordering temperature of Sm), at the Sm L_2 edge we observed a dipole resonance peak approximately 2 eV above the absorption edge for the (1 0 7.5) reflection. Similar energy scans were performed at the Sm L_3 edge and are shown in Fig. 5 (b). In addition to the dipole feature observed at the L_2 edge, a quadrupole feature appears approximately 6 eV below the Sm L_3 edge. Figure. 5(c) shows the energy scan through the Fe K-edge. Several features are observable in the energy spectrum: (a) Resonant features at and above E = 7.106 keV and (b) an energy independent non-resonant signal for energies below the resonant features. The non-resonant signal is about

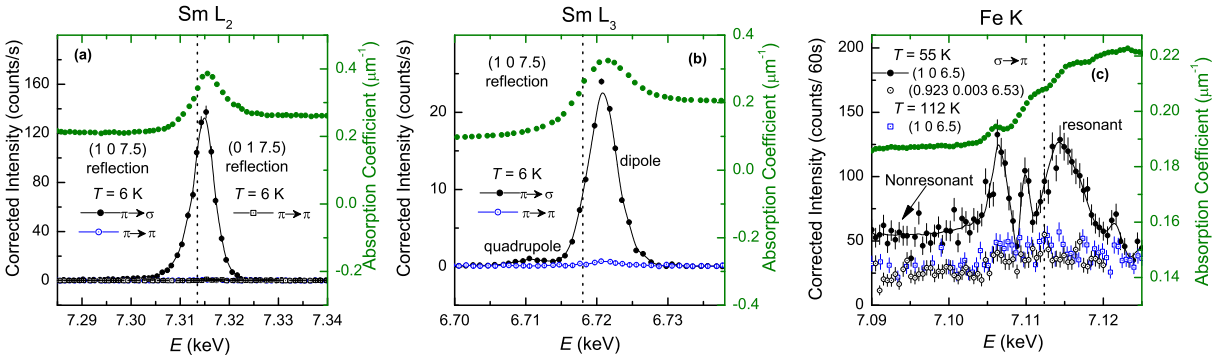


Figure 5: *(a,b) Energy scans of the (1 0 7.5) and (0 1 7.5) reflections and of the absorption coefficient at the Sm L_2 (left panel) and L_3 edges (middle panel). The dashed lines depict the Sm L_2 and L_3 absorption edges as determined from the inflection point of the absorption coefficient [22]. (c) Energy scans of the absorption coefficient and of the (1 0 6.5) reflection below (T = 55 K, filled circles) and above (T = 112 K, open squares) the magnetic ordering temperature of Fe moments and the measured background at T = 55 K away from the magnetic Bragg peak (open circles). The dashed line depicts the Fe K-edge. Both resonant and nonresonant signal can be clearly seen near the Fe K edge.*

a factor of 2.5 smaller than the resonant signal. To determine the polarization of the scattered x-rays, polarization analysis was performed using polarization analyzer which will discussed in more detail in section 4.4. Au (2 2 0) was used at Sm L_2 edge and Cu (2 2 0) was used for both the Sm L_3 and Fe K absorption edges as a polarization and energy analyzer to suppress the charge and fluorescence background relative to the magnetic scattering signal.

3 Poincaré Stokes Parameter:

To calculate the cross-section of magnetic scattering for any magnetic structure for an arbitrary incident polarization, it is convenient to introduce Poincaré representation for the the incident polarization and density matrix for the incident beam. In the co-ordinate system of non-resonant magnetic scattering amplitude the density matrix ρ can be written as:[10]

$$\rho = \frac{1}{2} \begin{pmatrix} 1+P_1 & P_2 - iP_3 \\ P_2 + iP_3 & 1-P_1 \end{pmatrix} \tag{13}$$

where $\vec{P} = (P_1,\ P_2,\ P_3)$ is the Poincaré-Stokes vector with components P_1, P_2, P_3. Taking $\hat{\epsilon}_\perp$ and $\hat{\epsilon}_{||}$ as two orthogonal unit vectors perpendicular and parallel to scattering plane (as in Fig. 3) we can write the incident electric field vector as:

$$\vec{E} = E_1\hat{\epsilon}_\perp + E_2\hat{\epsilon}_{||} \tag{14}$$

The components P_1, P_2, P_3 are defined as follows:

$$P_1 = \frac{|E_1|^2 - |E_2|^2}{|E_1|^2 + |E_2|^2} \quad (15)$$

$$P_2 = \frac{|E_1 + E_2|^2 - |E_1 - E_2|^2}{2(|E_1|^2 + |E_2|^2)} \quad (16)$$

$$P_3 = \frac{|E_1 + iE_2|^2 - |E_1 - iE_2|^2}{2(|E_1|^2 + |E_2|^2)} \quad (17)$$

P_1 = +1(-1) represent linear polarization along the $\hat{U}_2$ ($\hat{U}_3$) axis respectively. P_2 = +1(-1) represent linear polarization at an angle +45°(-45°) to the $\hat{U}_3$ axis, P_3 = +1(-1) represent left (right) circular polarization. If $\left|\vec{P}\right| = 1$, then the beam is completely polarized (100 %); $\left|\vec{P}\right| \leq 1$, the beam is partially polarized and for $\left|\vec{P}\right| = 0$, the beam is unpolarized. We should note that $\vec{P}$ is not a vector in real space, since it does not follow the transformation properties of a vector but can be thought of as vector in abstract space. If *M* represent the matrix for the scattering in the sample, then to calculate the density matrix and final polarization after scattering we need:

$$\vec{P} = tr(\sigma\vec{\rho}) \quad (18)$$

$$\rho' = M\rho M^\dagger \quad (19)$$

$$\frac{d\sigma}{d\omega} = \left(\frac{e^2}{mc^2}\right)^2 tr(\rho') \quad (20)$$

$$\vec{P'} = \frac{tr(\vec{\sigma}\rho')}{tr(\rho')} \quad (21)$$

where $\vec{\sigma}$ represents Pauli matrices. $\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $M^\dagger$ is the Hermitian conjugate of M.

4 Synchrotron instrumentation for the magnetic scattering:

For a magnetic scattering experiment we need a few basic components. (a) **The source**: From the discussion of the XRMS and NRXMS cross-sections, it is clear that for a successful magnetic scattering experiment one needs, polarized x-rays with high photon flux as well as tunability of the incident energy. Synchrotron radiation fulfills all the above criteria. (b) **The beamline**: This part prepares incoming x-ray from a synchrotron source for the experiment and consists of different optical elements such as monochromators (to change the wavelength of the incoming x-ray), focusing mirrors, slits (to define the beam cross-section), attenuators (to reduce the incident flux), and phase-plates (to change the incident polarization). All the optical elements are kept under vacuum and the beam is transported through evacuated tubes. (c) **The spectrometer**: For the observation of a diffraction peak, the sample has to be oriented within few tenths of a degree. This is done with a versatile diffractormeter with angular resolution better than 0.001°. A detector is attached with the 2θ arm and more often, before the detector a polarization analyzer is attached to analyze polarization of the scattered x-rays. (d) **Sample Environment**: Most of the cases, sample is placed in a cryostat to decrease the temperature of

the sample such that the sample is in a magnetically ordered state. This also allows measurement of magnetic order parameter as a function of temperature. Cryomagnets are used if one needs magnetic field for the experiment.

4.1 The source: synchrotron basics:

Here we will outline basic properties of synchrotron radiation relevant for the magnetic x-ray scattering. Basic electromagnetism tells us that an accelerating charge particle emits electromagnetic radiation. Synchrotron radiation occurs when an accelerating electrons or positrons follows a curved path in a relativistic speed. The electrons travel in a constant magnetic field which keep the electrons in a curved circular path.
Figure 6 shows the electron orbit in constant magnetic field. The direction of the radiation cone is in the direction of instantaneous velocity with opening angle ~$\gamma^{-1} = \frac{m_e c^2}{E_e}$ where $m_e c^2$ is the rest mass of an electron (= 0.511 MeV) and E_e is the energy of the circulating electrons. Typical energy E_e = 5 GeV of a third generation synchrotron radiation source gives the opening angle of the radiation cone $\gamma^{-1} \sim 0.1$ milli-radian. The synchrotron radiation is produced either in a bending magnet (BM) or some straight insertion devices are placed such as wigglers and undulators. The radiation coming out of a bending magnet is linearly polarized when viewed on axis. The beam is elliptically polarized above and below the axis with opposite helicities. The emitted radiation is very broad with emitted spectrum ranging from far-infra red to hard x-ray regime. However, the intensity of the emitted radiation falls of very rapidly for photon frequencies higher than $\gamma^3\omega_0$ where ω_0 is the angular frequency of an electron in the storage ring which is of the order of 10^6cycles/sec. Modern synchrotron radiation sources employ wiggler and undulators as insertion devices to increase the photon flux. These insertion devices consists of permanent magnets with opposite polls to create an alternating magnetic field perpendicular to the orbit of the storage ring. This alternating field leads to the sinusoidal movement of the charge particles withing the insertion devices. In an undulator or wiggler an important parameter characterizing the electron motion is the deflection parameter K given by: (see x-ray Data Booklet, http://xdb.lbl.gov/xdb-new.pdf)

$$K = \frac{eB_0}{2\pi m_e c} = 0.934\lambda_u[cm]B_0[T] \tag{22}$$

Where B_0 is the peak magnetic field and λ_u is the magnetic period. The maximum angular dispersion of the electrons in the orbit is given by $\delta = \frac{K}{\gamma}$. For an undulator, $K \lesssim 1$, and hence, the angular deviation is smaller than the opening of the radiation cone ($\sim \frac{1}{\gamma}$). Therefore, radiation from different poles adds coherently. The coherent addition of amplitude leads to monochromatic spectrum with odd harmonics on axis. Due to the finite number of magnetic poles the the radiation is quasi-monochromatic with a bandwidth of 0.1 %. The monochromaticity $\frac{\Delta\lambda}{\lambda}$ is inversely proportion to the number of polls and harmonics index number. For a wiggler, $K \gg 1(\sim 10)$, and the radiation from different polls adds incoherently.
The fundamental wavelength from an undulator radiation for an observation with an angle θ with the undulator axis is given by:

$$\lambda_1(\theta) = \frac{\lambda_u}{2\gamma^2}(1 + \frac{K^2}{2} + (\gamma\theta)^2) \tag{23}$$

Since $\lambda_u \sim 1$ cm and γ^{-2} is in the range 10^{-8}, the wavelength of the fundamental radiation lies in the Ångstrom range, i.e. in the x-ray regime. Furthermore, λ_1 can be varied by changing the

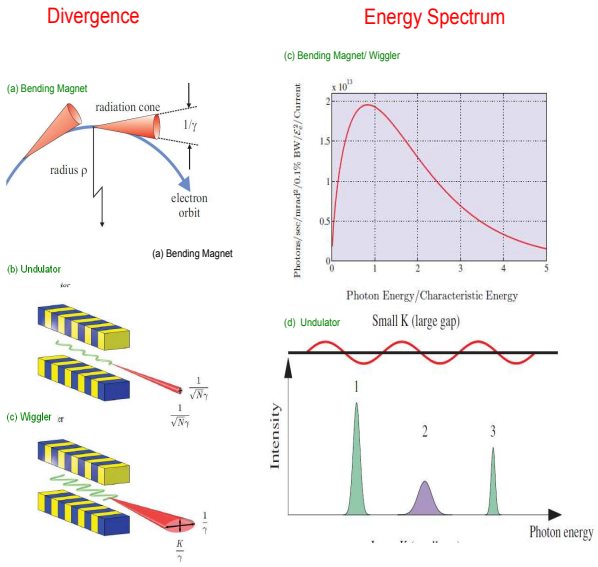


Figure 6: *(Left panel): Divergence properties synchrotron radiation of a Bending magnet, undulator and wiggler source. (Right panel): Energy spectrum from Bending magnet/wigglers and undulators. Figures taken from Ref. [23]*

magnetic field by varying the gap of the undulator (see. Eq. 22). The angular divergence of the undulator is substantially reduced compared to the bending magnet and the full width half maximum of the angular divergence is given by :[23]

$$\theta_{FWHM} \cong \frac{1}{\gamma}\sqrt{\frac{1+K^2/2}{nN}} \tag{24}$$

Where n is the harmonics index and N is the number of magnetic periods. This is much smaller than the angular divergence of the bending magnet which is $\sim \frac{1}{\gamma}$. The peak intensity of the undulator radiation scales with N^2 where as the central cone flux scales with N. Therefore, undulator radiation is highly collimated, linearly polarized in the plane, tunable with much improved flux which gives ideal condition for magnetic x-ray Scattering.

For the radiation from an wiggler, the radiated spectrum is same as that of a bending magnet with same filed strength, however with an intensity improved by a factor of 2N where N is the number of magnetic periods. The radiation is linearly polarized on-axis. However, off-axis polarization is still linear with different direction in compared with a bending magnet where off-axis polarization is elliptical. Divergence, energy spectrum and flux of the outcoming radiation from bending magnet (BM), undulators (ID) and wigglers has been compared in Fig. 6 and Fig. 7, respectively.

Worldwide, there exist four so-called "third generation synchrotron sources" for the hard x-ray regime: the European Synchrotron Radiation Facility ESRF in Grenoble, SPRING-8 in Japan, the Advanced Photon Source (APS) in Argonne, USA and the Petra III radiation source in Hasylab, DESY, Germany. All these facilities except ESRF have beamlines dedicated to magnetic scattering. The magnetic scattering beamline at ESRF has been recently closed after successful operation over years.

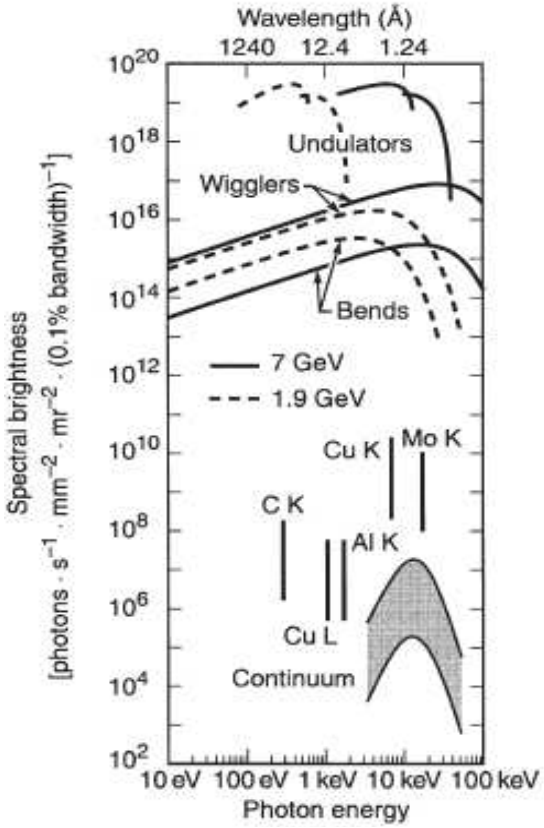


Figure 7: *Spectral brightness for the conventional laboratory x-ray tubes and several synchrotron radiation sources. The envelope of spectral brightness for today's third generation synchrotron facilities is somewhat higher than the curves above due to increased average beam current and decreased electron beam emmitance. Figure taken from x-ray data booklet.*

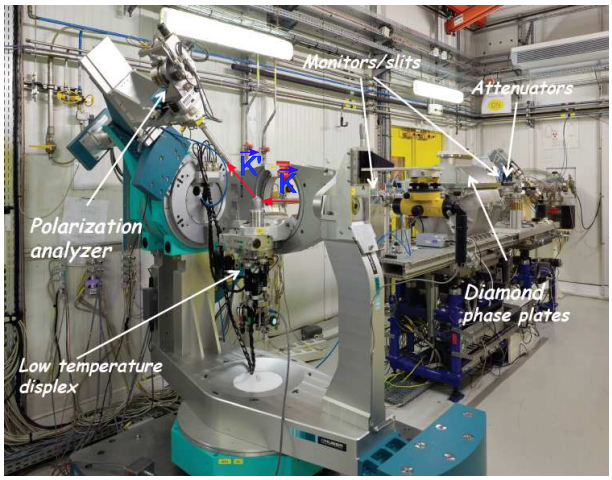


Figure 8: *Six-circle diffractometer at the ID20 beamline at ESRF, Grenoble.*

4.2 Beamlines:

The typical set-up of such a beamline is as follows: the beam from an undulator source is tailored by slit systems. Mirrors are used to suppress higher harmonics in the x-ray beam and/or to focus the beam onto the sample in the experimental hutch. The x-ray energy is selected by a double crystal Bragg monochromator. To handle the high heat load of several hundred W/mm^2, liquid nitrogen cooled silicon crystals are employed for monochromatization. Usually the Si (1 1 1) or Ge (1 1 1) reflection is chosen, since the second harmonic is largely suppressed for the Si crystal structure by the diamond glide planes. This beamline optics is situated in a so-called optics hutch with lead walls serving as biological radiation shielding. In the optics hutch, the x-ray beam is prepared with a desired properties and then enters the experimental hutch, where the actual scattering experiment is situated.

4.3 The Spectrometer:

Often, the six-circle diffractometers manufactured by Huber GmbH (http://www.xhuber.de) are used as a spectrometer. Figure 8 shows such a six circle diffractometer which was situated at the ID20 magnetic scattering bemline at ESRF. The diffractometer is equipped with a motorized analyzer stage and motorized translation stage. Motorized analyzer stage allows continuous rotation of the analyzer assembly around the scattered x-rays whereas motorized translation stage allows translation of the sample in three orthogonal directions. With this six circle diffractometer both the horizontal and vertical scattering planes can be utilized since the detector arm can be moved in the vertical direction as well as in the horizontal plane. Two types of detectors are generally used. The NaI photomultiplier tube and solid state detectors. NaI detectors has very low dark current ($\leq$0.1 counts/sec) and very high efficiency. However, it has very poor energy resolution.[24] On the other hand, solid state detectors have very high energy resolution which allows removal of fluorescence background with electronic discrimination.

4.4 Phase plate and the linear polarization analysis:

X-ray phase plates are used to manipulate polarization of the incident beam. One can change the direction of linear polarization using half wave plate and also, can convert linear polarization

into circular polarization using quarter-wave plates. Phase plates are based on the birefringence (anisotropy of refractive index) near Bragg-reflection. The component of transmitted electric field perpendicular (σ) and parallel (π) to the diffraction plane take a phase difference (Φ):[25]

$$\Phi = -\frac{\pi}{2}\frac{r_e^2\lambda^3 Re[F_h F_h^-]\sin(2\theta)t}{[\pi V^2]\Delta\Theta} \propto \frac{t}{\Delta\Theta} \tag{25}$$

where r_e is the classical electron radius, λ is the wavelength of the x-rays, t is the x-ray beam path in the crystal or the effective thickness, V is the volume of the unit cell, θ is the Bragg angle, F_h and $F_h^-$are the structure factors of the $(h\,k\,l)$ and $(\overline{h}\ \overline{k}\ \overline{l})$ reflections, respectively. The phase shift (Φ) is proportional to the effective thickness (t) of the crystal and inversely proportional to the angular offset ($\Delta\Theta$) from the Bragg peak. For a crystal with it's diffraction plane at an angle χ (see Fig. 4) and with a phase shift Φ, the Poincaré Stokes parameters are given by: [26]

$$P_1 = 1 + [\cos\Phi - 1]\sin^2 2\chi \tag{26}$$

$$P_2 = \sin^2\frac{\Phi}{2}\sin(4\chi) \tag{27}$$

$$P_3 = -\sin\Phi\sin(2\chi) \tag{28}$$

For $\Phi = \pm\frac{\pi}{2}$ and $\chi = 45^\circ$, the linear components P_1 and P_2 are zero whereas the circular component $P_3 = \pm 1$. This is known as Quarter Wave Plate (QWP) in optics which produces circularly polarized light from an incident linear polarization. The sign of the phase shift determines the helicity of the circular polarization, i. e. left or right circular polarization. If $\Phi = \pm n\pi\,(n = 0, \pm 1, \pm 2, ...)$, the the circular component is zero and the two linear components are given by: $P_1 = \cos(4\chi)$ and $P_2 = \sin(4\chi)$. This is known as Half Wave Plate (HWP). Variation of P_1 and P_3 as function of offset angle $\Delta\Theta$ for (1 1 1) diamond plate is shown in Fig. 9.

The polarization state of the scattered beam or the direct beam can be analyzed using a linear polarization analyzer (PA). A polarization analyzer is based on Thomson charge scattering from crystal such that the scattering angle (2θ) is close to 90°. The condition $2\theta = 90^\circ$ is fulfilled only approximately as a perfect mach of polarization analyzer crystal d spacing with the incident x-ray energy can not always be obtained. A list of different PA crystals in different energy range is listed in Table. 2. The intensity of the scattered radiation from an analyzer crystal can be written as:

$$I = \frac{I_0}{2}[1 + P_1\cos(2\eta) + P_2\sin(2\eta)] \tag{29}$$

where η is the rotation angle of the PA around the beam axis (see Fig. 4).

As an example, let us analyze the polarization of the direct beam with a Au(1 1 1) analyzer crystal. The energy is chosen such that $2\theta_P = 90^\circ$ condition is fulfilled. In this case, the polarization of the direct beam has been changed using a HWP by changing χ of the diamond phase plate. The rotation of incident light (ζ) is related to the rotation angle of diamond phase plate by : $\zeta = 2\chi$ (see Fig. 4). The integrated intensity of the scattered beam from the PA has been measured as function of η by rocking the analyzer crystal for a particular ζ and is plotted for different values of ζ in Fig. 10. The curves for a particular ζ has been fit with Eq. 29 and the values of P_1 and P_2 are extracted. These procedure is followed with different ζ and the

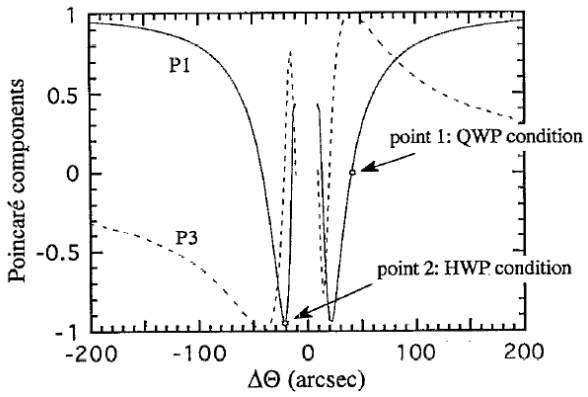


Figure 9: *Poincaré components (P_1 and P_3) as a function of offset angle $\Delta\Theta$ for a (1 1 1) diamond phase plate with 0.77 mm thickness in an (1 1 $\bar{1}$) asymmetric Laue geometry (beam path of 0.99 mm). Figure taken from Ref. [25]*

.

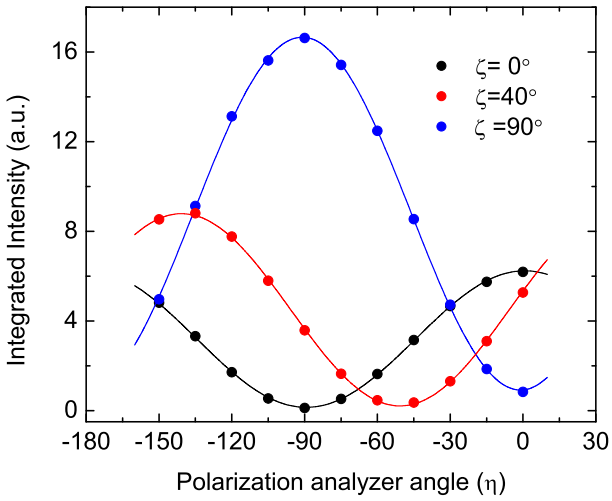


Figure 10: *Polarization analyzer angle (η) dependence of the scattered intensity for different incident linear polarization. The solid lines are the fits using Eq. 29 to extract Poincaré Stokes parameters.*

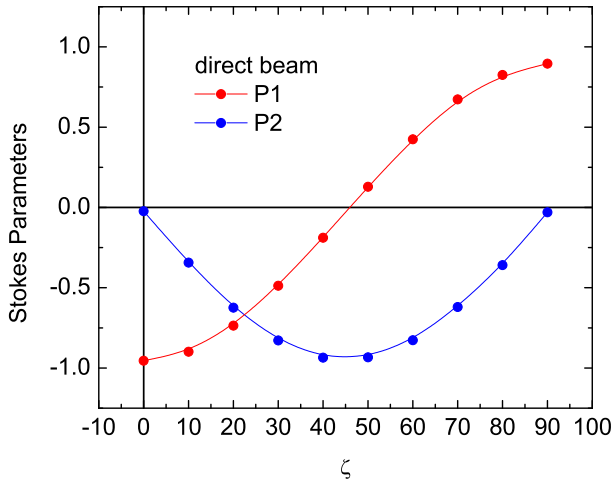


Figure 11: *Poincaré-Stokes parameters for direct-beam from a phase plate as a function of rotation of the incident light polarization.*

resulting values of P_1 and P_2 are plotted as a function of ζ in Fig. 11. The variation of P_1 and P_2 follows the theoretical prediction $P_1 = \cos(4\chi)$ and $P_2 = \sin(4\chi)$. Note that the degree of circular polarization can not be determined in this way. However an upper limit can be placed since $P_3^2 \leq 1 - P_1^2 - P_2^2$.

It is also possible to determine P_1 and P_2 by collecting integrated intensities for a pair of η values since P_1 and P_2 can be written from Eq. 29 as $P_1 = [I(0°) - I(90°)]/[I(0°) + I(90°)]$ and $P_2 = [I(+45°) - I(-45°)]/[I(+45°) + I(-45°)]$. However, it is recommended to measure integrated intensity at different positions of η with step size of minimum 30° between 0° and 180° to estimate the systematic error [27]. Furthermore, the integrated intensities should be measured by rocking the analyzer crystal to avoid the artifacts from the variable resolution function as a function of η.

There are several benefits of using polarization analyzer. PA analyzer reduces the background signal from fluorescence and diffuse signal. It also reduces the remaining higher harmonics.

5 Applications of the magnetic x-ray scattering:

X-ray magnetic scattering have many applications ranging from the study of the thin film magnetic system to the magnetism of bulk. Here few examples are selected to illustrate the power and versatility of the resonant and nonresonant magnetic scattering.

5.1 Determination of moment direction:

For a magnetic structure determination a basic and first step is to determine the direction of magnetic moments within the unit cell. In different ways this can be done, viz. (a) azimuthal rotation (b) Q-dependent measurements (c) by analyzing Poincaré Stokes parameters of the scattered beam. We will discuss all the procedures in the following sections with examples selected from different compounds with present relevance.

Table 2: *A list of analyzer crystals which are commonly used for polarization analysis. The photon energy for $2\theta_{PA} = 90^\circ$ is presented in columns 3 and 6. PG stands for Pyrolytic graphite.*

Crystal	(H K L)	*E* (90°) [keV]	Crystal	(H K L)	*E* (90°) [keV]
Au	(1 1 1)	3.73	Pt	(2 2 0)	7.74
Al	(1 1 1)	3.75	PG	(0 0 6)	7.84
LiF	(2 0 0)	4.35	Ge	(3 3 3)	8.05
Cu	(2 0 0)	4.85	Si	(3 3 3)	8.39
PG	(0 0 4)	5.22	Cu	(2 2 2)	8.41
Mo	(2 0 0)	5.57	LiF	(4 0 0)	8.71
Al	(2 2 0)	6.12	Pt	(4 0 0)	8.94
LiF	(2 2 0)	6.16	Pd	(4 0 0)	9.01
Cu	(2 2 0)	6.86	Cu	(4 0 0)	9.70
Au	(2 2 2)	7.44	PG	(0 0 8)	10.48
Al	(2 2 2)	7.49	Au	(3 3 3)	11.16

Azimuthal angle dependence:

$Gd_5(Si_xGe_{1-x})_4$ have received attention due to their magnetocaloric, magnetostrictive and magnetoresistive properties [28, 29, 30, 31]. Tan *et al.* [32] have undertaken determination of magnetic structure of the Gd moments in Gd_5Ge_4 using XRMS at the Gd L_{II} edge. Since naturally occurring Gd is strongly neutron absorbing, XRMS is a feasible option to determine the magnetic structure of this compound.

Magnetization measurement indicate that the Gd_5Ge_4orders antiferromagnetically below T_N = 127 K. Also, the measurements of magnetic susceptibility along three crystallographic directions indicate that the moment is along the *c* direction. Therefore, the crystal was mounted with *b* axis vertical and the *c* and *a* axis in the horizontal plane. The magnetic (0 3 0) reflection was accessed. In this geometry the crystal can be rotated around the $\vec{Q} = (0\,3\,0)$ such that the angle ψ between the *c* axis and the scattering plane can be continuously changed from 0° to 180°. The rotation of crystal around a scattering vector without changing value of Q vector is know as azimuthal rotation. A cartoon of azimuthal rotation is shown in Fig. 12.

By rotating the crystal around the *b* axis either *b*-*c* or *a*-*b* can be brought into the scattering plane. The integrated intensity of the (0 3 0) magnetic peak was collected as a function of ψ and was normalized to the charge (0 4 0) reflection for every azimuth angle to reduce the systematic error and is shown in Fig. 13. The dipole scattering intensity is sensitive only to the moment in the scattering plane with a cross section $f \propto \hat{k'}\cdot\hat{\mu}$ ($\hat{k'}$ and $\hat{\mu}$ are the wave vectors of the scattered photon and the magnetic moment respectively.) With $\psi = 90^\circ$ (a-b in the scattering plane), the integrated intensity is close to zero. Whereas when $\psi = 90^\circ$ or 180° (*b*-*c* in the scattering plane), the integrated intensity is maximum. Therefore, only *c* component of the magnetic moment contributes to the resonant scattering. Let us assume that the moments are at an angle ψ_c with respect to the *c* axis. Then the angle between the magnetic moment and the scattering plane for an azimuth angle ψ (defined as the angle between *c* axis and the scattering plane) is equal to $(\psi - \psi_c)$. Therefore, $f \propto \cos(\psi - \psi_c)$ and the integrated intensity can be written as $I = A\cos^2(\psi - \psi_c)$ where *A* is an arbitrary scaling factor. The solid line in Fig. 13

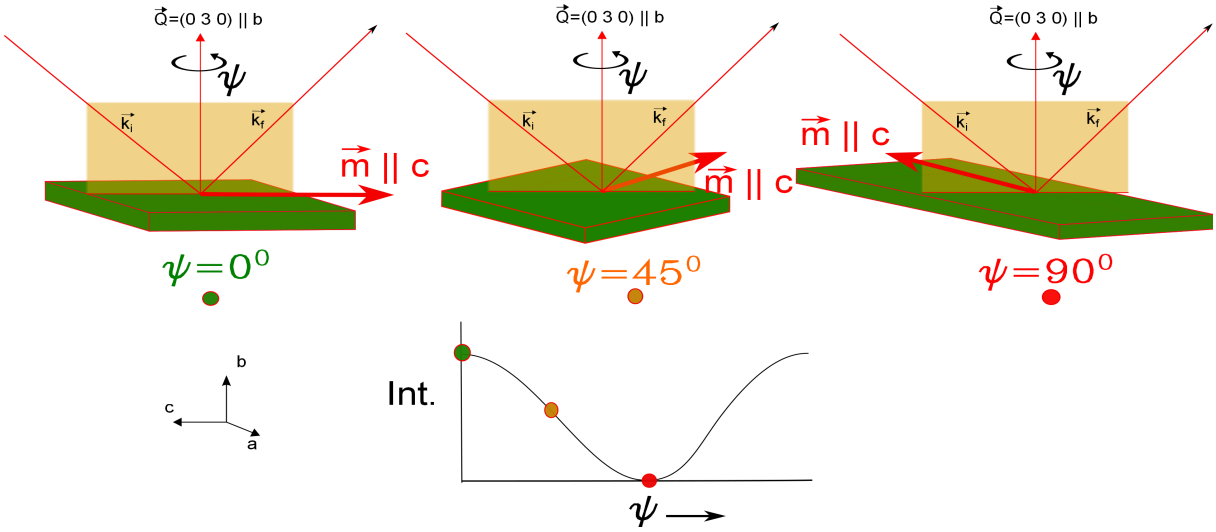


Figure 12: *Schematic representation of the azimuthal rotation. The ψ angle is defined as the angle between the crystallographic* c *axis and the scattering plane. When* c *axis lies within the scattering plane, ψ is defined as zero. When ψ = 90°, the* c *axis is perpendicular to the scattering plane. Expected variation of the dipole resonance intensity for the magnetic moment along the* c *direction.*

represents a fit with the above equation with $\psi_c = 1.9° \pm 1.8°$. Therefore, the magnetic intensity to the (0 3 0) reflection is sensitive only to the moment along *c* direction and indicate that either there is no magnetic component along *a* or *b* axis or the intensity of the (0 3 0) magnetic peak is not sensitive to either the *a* or *b* magnetic components due to cancellations arising from the symmetry of the magnetic order. Later, it was verified by a detailed Q-dependent measurements that the magnetic moments are indeed along the *c* direction.

One may note that the integrated intensity at $\psi = 0°$ deviates significantly from the calculated curve. This is due to the presence of multiple charge scattering. [33] Multiple charge scattering is particularly a problem for commensurate magnetic structures. Multiple charge scattering arises due to two or more successive Thompson scattering events. It occurs when another lattice point $\vec{Q}_2$ (secondary reflection) intercept the Ewald sphere other than the lattice point of interest $\vec{Q}_1$ (primary reflection). A third reflection $\vec{Q}_3 = \vec{Q}_1 - \vec{Q}_2$ is required to bring the secondary reflection into the direction of main reflection. In this case, not only $\vec{Q}_1$ but also $\vec{Q}_2 + \vec{Q}_3$ contribute to the observed intensity. Multiple diffraction depends on several factors such as incident energy of the x-rays, azimuth angle and crystal symmetry, lattice constants etc. For a resonant magnetic scattering experiment resonance occurs at particular energy or within a very small energy range. Therefore, one can reduce the multiple charge scattering background by varying the azimuth angle (rotation around the primary reflection $\vec{Q}_1$) such that $\vec{Q}_2$ is out of the Bragg condition. Typically a very small rotation ($\leq 1°$) is needed. The problem of multiple charge scattering becomes severe for high energy magnetic diffraction since size of the Ewald sphere is larger and correspondingly, more and more secondary reflection can intercept the Ewald sphere. In practice, energy scans around the resonant energy as a function of azimuth angle is performed to select multiple charge scattering free region. Fig. 14 (a) shows contour map of intensity in dependence on energy and azimuth angle for the (5 0 0) reflection. The multiple scattering contribution at the resonant energy can be minimized through a judicious choice of azimuth angle as shown in Fig. 14 (b), where the resonant scattering is well separated

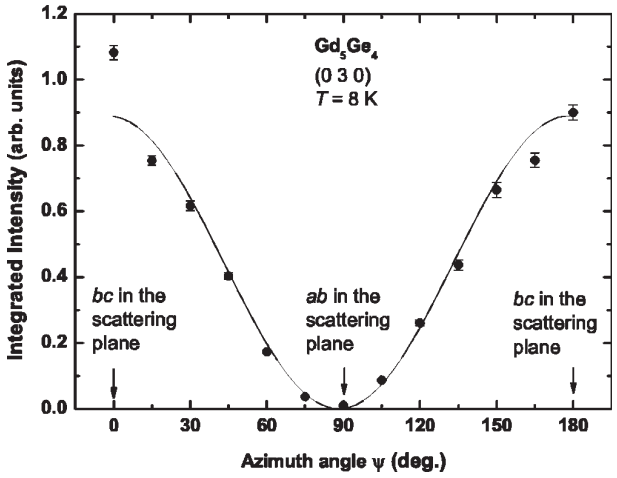


Figure 13: *The integrated intensity of the (0 3 0) magnetic peak normalized by the (0 4 0) charge peak at T = 8 K. The solid curve represents the variation expected for magnetic moments along the c axis. The Figure has been taken from Ref.[32].*

from the multiple scattering.

Q-dependence:

The conventional way to determine the moment direction is to measure the scattered intensity as a function of scattering angle which is commonly known as "Q dependent" measurement in the scattering community. There are many examples of using this procedure to determine the moment direction and magnetic structure [15, 14]. Here we will present the case of SmFeAsO [34]. SmFeAsO is the parent compound for the recently discovered superconducting $SmFeAsO_{1-x}F_x$ compounds where the highest T_c of 55 K has been observed. [35, 36]

SmFeAsO crystallizes in the tetragonal $P4/nmm$ space group and undergoes a structural phase to an orthorhombic crystal structure $Cmme$ below T_S = 140±1 K. Below T_{N1} = 110 K, a magnetic signal was observed at the reciprocal lattice points characterized by the propagation vector $(1\,0\,\frac{1}{2})$ when the x-ray energy was tuned through the Sm L_2 and Fe K-edges, indicating the onset of Sm and Fe magnetic order, respectively.

For the determination of the magnetic structure in the temperature range $5\,K \leq T \leq 110\,K$, one has to look into the details of the magnetic structures allowed by the space group symmetry and the propagation vector $(1\,0\,\frac{1}{2})$. Six independent magnetic representations (MRs) are possible [38]. All the MRs along with the calculated intensities for different polarization geometries are listed in Table 3 . For a second-order phase transition, Landau theory predicts that only one of the six above mentioned MRs is realized at the phase transition [38]. We note that the $\pi \rightarrow \pi'$ scattering geometry is sensitive only to the moment perpendicular to the scattering plane for the dipole resonance . Since, no magnetic signal was observed at the (0 1 7.5) (sensitive to Γ_1 and Γ_8) and (1 0 7.5) (sensitive to Γ_2 and Γ_7) reflections in the $\pi \rightarrow \pi'$ scattering channel at the Sm L_2 edge, one can exclude the moment in the a and b directions and hence, the MRs Γ_1, Γ_8, Γ_2 and Γ_7. To differentiate between the MRs Γ_4 and Γ_5 (moment along the c direction), the integrated intensities for a series of $(1\,0\,\frac{L}{2})$ reflections were measured, see Fig. 15(a).

To differentiate between these two MRs, angular dependence of the magnetic scattering has

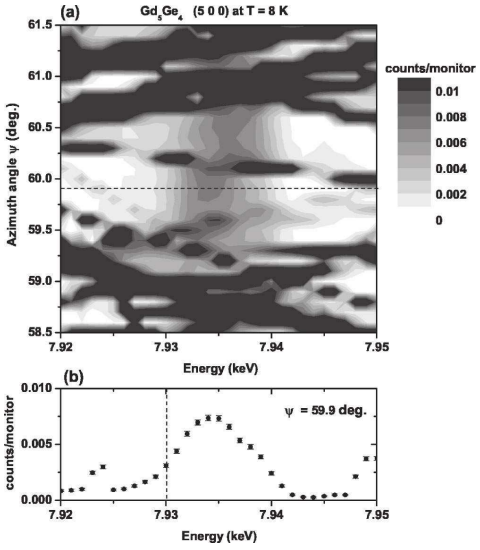


Figure 14: *(a) Contour map of the intensity as a function of energy and azimuth angle ψ at the (5 0 0) position and T = 8 K. Discontinuities in the bands of multiple scattering across the energy range are artifacts of the scanning process and (b) single energy scan at the azimuth angle* c *= 59.9°, which is depicted as a horizontal dashed line in (a). In (b), the vertical dashed line represents the position of the Gd L_{II} absorption edge. Figure adapted from Ref. [32].*

been calculated as outlined below. The intensity for a particular reflection can be written as:

$$I = SAL|F_m|^2 \tag{30}$$

where *S* is arbitrary scaling factor, $A = \frac{\sin(\theta+\alpha)}{\sin\theta\cos\alpha}$ is the absorption correction, $L = \frac{1}{\sin 2\theta}$ is the Lorentz factor. $|F_m|$ is the modulus of the magnetic structure factor. The magnetic structure factor F_m for the (*h k l*) reflections can be written as:

$$F_m = \sum_j f_j e^{2\pi i(hx_j+ky_j+lz_j)} \tag{31}$$

The summation is over all the magnetic atoms in the unit cell. f_j is the resonant/non-resonant magnetic scattering amplitude. The angular dependence of the magnetic structure factor comes from the magnetic scattering amplitude (f_j). For dipole resonance and for the $\pi \rightarrow \sigma$ geometry $f \propto \hat{k_i}\cdot\hat{\mu}$ where $\hat{k_i}$ and $\hat{\mu}$ are the wave vectors of the incoming photons and the magnetic moment, respectively. Here, α is the angle that the scattering vector $\vec{Q}$ (= $\vec{k_f}$-$\vec{k_i}$) makes with the crystallographic *c* direction perpendicular to the surface of the sample and, θ is half of the scattering angle. α is positive/negative for larger/smaller angles for the outgoing beam with respect to the sample surface. For the dipole resonance, and for the reflections of the type $(1\,0\,\frac{L}{2})$, $|F_m|^2$ is proportional to $\sin^2(2\pi zL)\sin^2(\theta+\alpha)$ and $\cos^2(2\pi zL)\sin^2(\theta+\alpha)$ for the Γ_4 and Γ_5 MRs, respectively. $z = 0.137$ is atomic position of Sm moments within the unit cell [37]. While $\sin^2(2\pi zL)/\cos^2(2\pi zL)$ term comes from the relative orientation of the magnetic moment within the magnetic unit cell, the term $\sin^2(\theta+\alpha)$ comes from the dot product between $\vec{k_i}$ and $\hat{\mu}$ [$(90^\circ - \theta - \alpha)$ is the angle between $\hat{k_i}$ and $\hat{\mu}$].

We note, that there is only one free parameter for the dipole intensity (see Eq. 30), namely the arbitrary scaling factor S. Figure 15(a) shows a fit to the observed intensities for the two above mentioned MRs. Since the model calculation with the magnetic moment in the Γ_5 MR closely

Table 3: *Basis vectors for the space group* $Cmme$ *with* $\mathbf{k}_{17} = (0,\ 1,\ .5)$. *The decomposition of the magnetic representation for the* Sm *site* $(0,\ .25,\ .137)$ *is* $\Gamma_{Mag} = 1\Gamma_1^1 + 1\Gamma_2^1 + 0\Gamma_3^1 + 1\Gamma_4^1 + 1\Gamma_5^1 + 0\Gamma_6^1 + 1\Gamma_7^1 + 1\Gamma_8^1$. *The atoms of the nonprimitive basis are defined according to 1:* $(0,\ .25,\ .137)$, *2:* $(0,\ .75,\ .863)$. *Lattice parameters of the orthorhombic crystal at 100 K [37]:* $a = 5.5732\,\mathring{A}$, $b = 5.5611\,\mathring{A}$, $c = 8.4714\,\mathring{A}$.

IR	Atom	BV components			Magnetic Intensity			
		m_a	m_b	m_c	$(h\,0\,\frac{l}{2})$		$(0\,k\,\frac{l}{2})$	
					$\pi \to \sigma$	$\pi \to \pi$	$\pi \to \sigma$	$\pi \to \pi$
Γ_1	1	1	0	0	Yes	No	No	Yes
	2	-1	0	0				
Γ_2	1	0	1	0	No	Yes	Yes	No
	2	0	1	0				
Γ_4	1	0	0	1	Yes	No	Yes	No
	2	0	0	1				
Γ_5	1	0	0	1	Yes	No	Yes	No
	2	0	0	-1				
Γ_7	1	0	1	0	No	Yes	Yes	No
	2	0	-1	0				
Γ_8	1	1	0	0	Yes	No	No	Yes
	2	1	0	0				

agrees with the observed intensity, we conclude that the magnetic Sm moments are arranged according to the MR Γ_5. Using the Q dependence of the non-resonant scattering, the magnetic structure of the Fe moments can be determined as well. A combined magnetic structure is shown in Fig. 15 (b).

Full polarization analysis using Stokes parameters:

Here we will discuss, how to use linear polarization analysis to determine the moment direction. There are several advantages of using Stokes parameter formalism over the conventional azimuthal scan to determine the moment direction. During the azimuthal scan the sample is rotated and during this rotation the beam can illuminate different parts of the sample if the sample is not precisely centered as it is often the case. For a multi-grain sample, this can increase the systematic error in the measured intensity as a function of azimuth angle. For a commensurate magnetic structure with the presence of multiple charge scattering, in principle, one has to minimize multiple charge scattering at each values of azimuth angle. This procedure might be time consuming. In contrast, one can obtain the same information by changing the incident polarization. The beam is fixed on the sample, and hence the previous problems can be easily avoided. Here we will outline the procedure for full polarization analysis used for the multiferroic compound $TbMn_2O_5$.

$TbMn_2O_5$ is a well known multiferroic where ferroelectric and magnetic order occurs in the same phase. In this compound a complete reversal of ferroelctric polarization has been observed in a small applied magnetic field (~2 T). [39] The mechanism driving multiferroicity in this compound is not fully understood. From several studies, it is clear that rare-earth plays an important role in the magnetoelectric coupling. It was not possible to refine the magnetic struc-

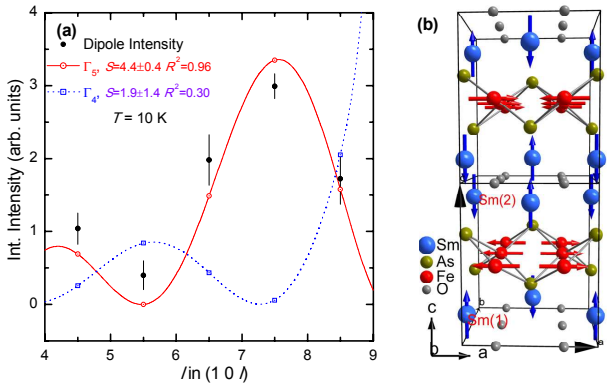


Figure 15: *(a) l dependence of the integrated intensity at the Sm L_2 edge along with the fits for the (1 0 $\frac{l}{2}$) reflections. Open symbols are the calculated intensities. Lines serve as guides to the eye. (b) Proposed magnetic structure in the temperature range 5 K $\leq$ T $\leq$ 110 K. Figure taken from Ref. [34].*

ture of Tb moments in this compound due to a very large no of free parameters in the neutron scattering refinement and the presence of another dominant magnetic subsystem, namely Mn. Resonant magnetic scattering offers an alternative to to study magnetism of Tb moments alone due it's elemental specificity.

$TbMn_2O_5$ shows a series of magnetic transitions. At T_N = 43 K, the manganese moments orders in an incommensurate structure with magnetic propagation vector (δ, 0, τ) with $\delta \approx 0.5$ and $\tau \approx 0.3$. At a slightly lower temperature, T_{FE} =38 K, the system enters into a ferroelectric phase with ferroelectric polarization along the *b* direction. Between $33 \geq T \geq 24$ K, the system locks into a commensurate phase with $\delta = 0.5$ and $\tau = 0.32$. [40, 41] From a full polarization analysis Johnson *et al.* [42] have shown that in the commensurate phase the Tb 5*d* bands are polarized by the manganese subsystem.

A outline to determine the magnetic structure using full polarization analysis will be provided here. At an atomic site (j) the magnetic moment vector can be written as:

$$\vec{m}(j) = m_1\hat{z}_1 + m_2\hat{z}_2 + m_1\hat{z}_3 \tag{32}$$

where m_1, m_2, m_3 are the components of magnetic moment along the directions of $\hat{z}_1$, $\hat{z}_2$, $\hat{z}_3$ respectively (see Fig. 3). The scattering matrix can be calculated for dipole resonance as follows:

$$\begin{aligned} A(\sigma \rightarrow \pi') &= \sum_{j}^{unit-cell} [m_1(j)\cos\theta + m_3(j)\sin\theta]e^{i\vec{q}\cdot\vec{r}_j} \\ &= F(m_1)\cos\theta + G(m_3)\sin\theta \end{aligned} \tag{33}$$

where F and G are two different functions of m_1 and m_3, respectively. In a similar way $\sigma \rightarrow \sigma'$ and $\pi \rightarrow \pi'$ components can be evaluated. Therefore, the scattering matrix can be calculated for dipole resonance can be written as:

$$\begin{aligned} \langle M_m \rangle &= \begin{pmatrix} A_{\sigma\rightarrow\sigma'} & A_{\sigma\rightarrow\pi'} \\ A_{\pi\rightarrow\sigma'} & A_{\pi\rightarrow\pi'} \end{pmatrix} \\ &= \begin{pmatrix} 0 & F(m_1)\cos\theta + G(m_3)\sin\theta \\ -F(m_1)\cos\theta + G(m_3)\sin\theta & L(m_2)\sin 2\theta \end{pmatrix} \end{aligned} \tag{34}$$

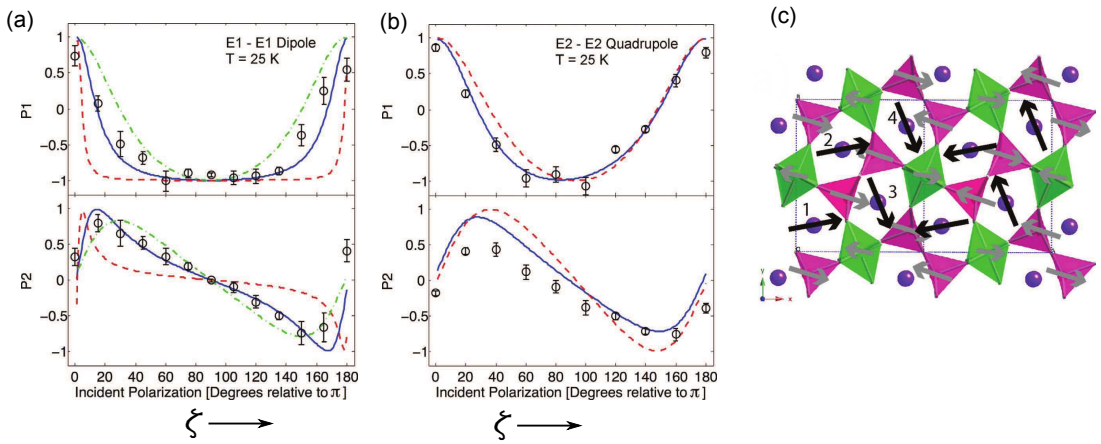


Figure 16: *(a) A plot of the measured Poincaré-Stokes parameters P1 and P2 as a function of the incident x-ray polarization of the commensurate reflection (4+δ, 4, 0-τ) at the E1-E1 (dipole) energy resonance at 25 K. Simulations of the Mn^{4+} and Mn^{3+} magnetic structures as refined by Blake et al. [Ref. [41]] are shown as blue solid and red dashed lines, respectively. A simulation of the terbium magnetic structure scattering at the E1-E1 transition in this phase, as refined in Fig. (b) at the E2-E2 transition, is shown by the green dash-dotted line. (b) A plot of the measured Poincaré-Stokes parameters P1 and P2 as a function of incident x-ray polarization of the commensurate (4+δ, 4, 0-τ) reflection at the E2-E2 (quadrupole) energy resonance at 25 K. (c) Diagrammatic illustration of the refined terbium ion magnetic-moment directions (black arrows) in a the CM phase. Figures taken from Ref. [42] .*

where $L(m_2)$ is a different function of θ. The density matrix of incident linear polarized light with a rotation ζ can be written as:

$$\rho = \frac{1}{2}\begin{pmatrix} 1+\cos 2\zeta & \sin 2\zeta \\ \sin 2\zeta & 1-\cos 2\zeta \end{pmatrix} \tag{35}$$

The final polarization of the scattered beam from the sample can be calculated as follows

$$P_1' = \frac{tr(\sigma_1\rho')}{tr(\rho')} = f(m_1, m_2, m_3, \zeta) \tag{36}$$

$$P_2' = \frac{tr(\sigma_2\rho')}{tr(\rho')} = g(m_1, m_2, m_3, \zeta) \tag{37}$$

where $\rho' = M\rho M^{\dagger}$. P_1' and P_2' can be measured experimentally using a linear polarization analyzer for a particular magnetic reflection as function of ζ and can be refined simultaneously with equation 36 and 37 to obtain the values of the three magnetic moment components. For $TbMn_2O_5$, P_1' and P_2' were measured using a Au(2 2 2) analyzer crystal (see the procedure mentioned in section 4.4) for the commensurate (4+δ, 4, 0-τ) reflection at the dipole and the quadrupole resonance at the Tb L_{III} absorption edge and is shown in Fig. 16(a) and (b) respectively.

Here we recall that the dipole resonance is sensitive to the polarization of the 5*d* band whereas sensitivity to 4*f* magnetism comes from the quadrupole resonance. Fig. 16(a) shows three simulations: (1) The solid blue line is the simulation assuming that the Tb moments are arranged according to the AFM superexchange interaction with nearest Mn^{4+} moments.[43] (2) The red dashed line is the simulation assuming the interaction between Tb and Mn^{3+} moments and (3) the green dashed line is the simulation according to the interaction with the Tb 4*f* moments,

the direction of which is refined in Fig. 16(b). This clearly established that the Tb 5*d* band are polarized by the close proximity of the Mn^{4+}spin density wave, as opposed to Mn^{3+} spin density wave. A reduced χ^2 value of 2.8 of the Mn^{4+} simulation, when compared to the value of 5.8 of the Tb 4*f* band simulation also shows that the Tb 5*d* band is predominantly polarized by the Mn^{4+} 3d band; however, one clearly cannot exclude the interaction with the terbium 4*f* spin configuration.
The magnetic moment direction of the Tb moments was also refined from Fig. 16(b) with moment directions on the 1 and 2 sites were refined to be $10.5\pm2.6^\circ$ in the *a*-*b* plane relative to the *a*-axis and $0.2\pm0.1^\circ$out of the plane and on the Tb 3 and 4 sites $292.5\pm2.0^\circ$in the *a*-*b* plane relative to the *a*-axis and $0.2\pm0.1^\circ$ out of the plane (see Fig.16(c)).
Here we note that the resonant scattering is not only element specific but also band selective. In this case, the Tb 4*f* and 5*d* bands were separately probed by tuning the x-ray energy to the quadrupole and dipole resonances, respectively. The full polarization analysis provides only the moment directions. To determine the absolute value of the magnetic moment, comparison of the nonresonant magnetic intensity with charge intensity is needed and was not performed in this case. One disadvantage of using phase plate is that the incident flux is reduced by at least a factor of ten due to absorption in the diamond crystal. Therefore, it is difficult to do full polarization analysis for small magnetic moment system where the magnetic scattering signal is small.

5.2 Separation of the spin and orbital angular momentum:

We have already discussed during the formulation of nonresonant magnetic scattering cross-section that it is possible to separate spin and orbital angular momentum component using specific polarization properties of the cross section. Here we will discuss the results of Ref. [44] in light of the separation of spin and orbital angular momentum although there exist many other examples [10, 45, 46]. Monooxides of 3*d* transition metals such as MnO, FeO, CoO and NiO are the test samples for the band theory models due to it's simplistic crystal and magnetic structures. Their insulating behavior contradicts simple electronic band models in which the oxygen *p* states are fully occupied while the metal *s* states are empty and the metal *d* states are partially occupied. To resolve this contradiction, two explanations have been proposed: the Mott-insulator concept [47] and band calculations based on local-spin-density approximation that take into account the antiferromagnetic order.[48] The latter approach implies that orbital moment plays a role in CoO and FeO, while such a contribution is not required in NiO. [49] Therefore, the determination of the orbital-moment contribution to the magnetization in the ordered state of NiO is a valuable piece of information towards a better understanding of the electronic and magnetic properties of these compounds.
NiO has the NaCl fcc structure with a = 4.177 Å at room temperature. The ground-state configuration of the Ni^{2+} ion has a $3d^8$ configuration. Below T_N =523 K, NiO orders in the type-II antiferromagnetic structure [50] where ferromagnetic planes are stacked antiferromagnetically along the [1 1 1] axes with their magnetic moments aligned in the [1 1 1] planes along one of the [1 1 $\overline{2}$] directions [51] . The experiment was performed at the ID20 magnetic scattering beamline at the ESRF. The sample was mounted on a four circle diffractometer. A [1 1 1]-axis normal to the surface was carefully oriented along the Φ axis of the four-circle diffractometer. The nonresonant scattering was measured at E = 7.84 keV to match the $2\theta_p$= $90^\circ$of the PG(0 0 6) analyzer crystal. The magnetic signal was measured at both the $\sigma \rightarrow \sigma'$ and $\sigma \rightarrow \pi'$ channels. Before, going into details, it is worth to discuss the magnetic domains present in the crystal.

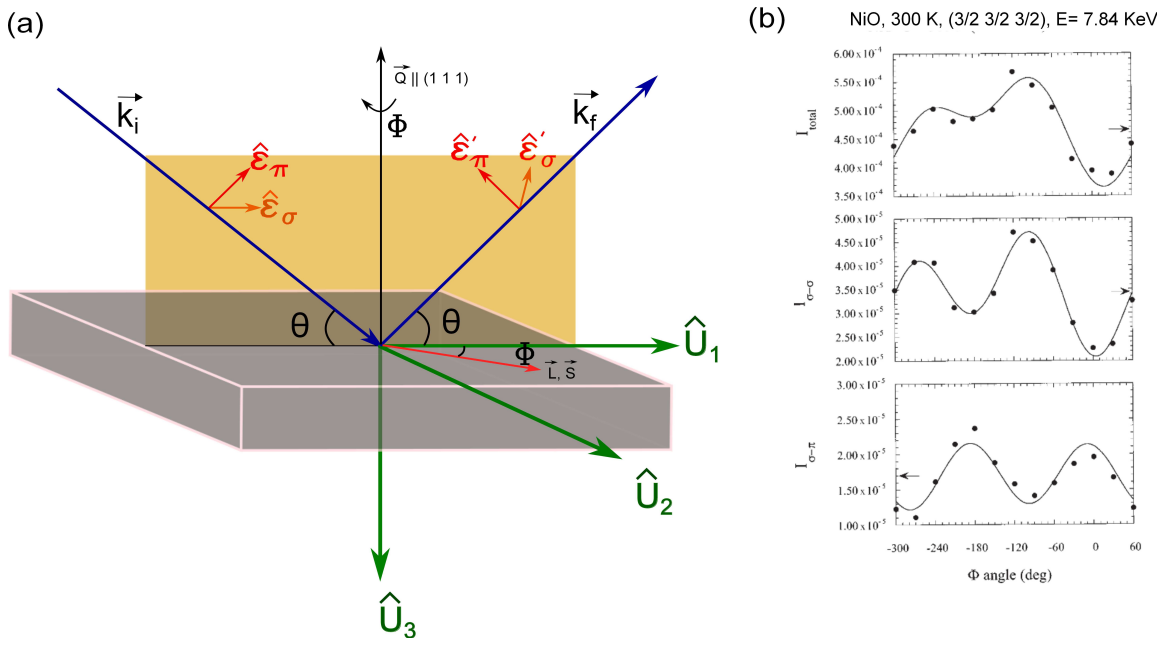


Figure 17: *(a) The scattering geometry for the NiO experiment. (b) Normalized integrated intensities of the* $(\frac{3}{2}\ \frac{3}{2}\ \frac{3}{2})$ *reflection as a function of azimuth angle* Φ *at 7.84 keV. Intensities are given in (radians*detector counts/monitor counts). Errors bars are smaller than the dot size. The two polarized components,* $I_{\sigma\sigma'}$ *and* $I_{\pi\pi'}$*, are out of phase by* $\frac{\pi}{2}$*. The total intensity is the sum of the two components corrected for the reflectivity of the PG(006) analyzer. Arrows indicate the* Φ *averaged values that are used to extract* $L(\vec{Q})$ *and* $S(\vec{Q})$*.*

When the system is cooled from the paramagnetic cubic phase to orthorhombic AFM phase, there exist 4 symmetry equivalent K domains. The propagation vector $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ can align in any of the 4 symmetry equivalent directions [1 1 1], [$\bar{1}$ 1 1], [1 $\bar{1}$ 1], [1 1 $\bar{1}$] and give rise to four K domains. In addition, the magnetic moments are aligned in the [1 1 1] planes along one of the [1 1 2] directions. Since in the cubic state, the [1 1 1] axis is a three-fold symmetry axis, there exist three equivalent possible spin directions perpendicular to the [1 1 1] axis, the so-called S-domains. Depending on crystal faults (surfaces, small angle grain boundaries, impurity atoms etc.), but also on random processes, an arrangement of these K- and S-domains develops in the sample in the low temperature phase. During the experiment it appears that the sample is populated by a single K domain with propagation vector parallel to the surface normal, at least in the near-surface region of 40 μm probed with 7.84 keV x-rays. When measuring the scattering intensities from a given K domain, all contributions from the associated S domains add incoherently. By rotating the sample about the surface normal (angle Φ in Fig. 17 (a)), we could study the S domain distribution within the [1 1 1] K domain. Figure 17 (b) shows the Φ dependence of $I_{\sigma\sigma'}$ and $I_{\pi\pi'}$ at the $(\frac{3}{2}\ \frac{3}{2}\ \frac{3}{2})$ position. The two intensities exhibit a modulation of period π characteristic of the S-domain distribution. A modulation with a period 2π would indicate that the footprint of the beam is moving across the surface of the sample during the Φ rotation. In a given S domain, the magnetic scattering amplitudes vary in a simple manner with the Φ angle: From the non resonant magnetic scattering cross-section we can write:

$$M(\sigma \rightarrow \sigma') = \sin(2\theta)S_2 = \sin(2\theta)\sin\Phi S(\vec{Q}) \quad (38)$$

$$M(\sigma \rightarrow \pi') = -\sin(2\theta)\sin\theta[L_1 + S_1] \quad (39)$$

$$= -\sin(2\theta)\sin\theta[\cos\Phi S(\vec{Q}) + \cos(\Phi + \Phi_0)L(\vec{Q})] \quad (40)$$

where an angular Φ_0 offset is allowed between *S* and *L*. The origin of Φ is taken with the spin direction in the scattering plane. In a multi-*S*-domain sample, scattered intensities are combinations of $\sin^2\Phi$ and $\cos^2\Phi$. Fig. 17(b) shows that the intensity at the $\sigma \rightarrow \sigma'$ and $\sigma \rightarrow \pi'$ are exactly $\frac{\pi}{2}$ phase shifted implying that Φ_0 = 0 which is in agreement with the collinear arrangement of the spin and orbital angular momentum. Taking the ratio of intensities ($I \propto M^2$) using Eqs. 38 and 40, in both geometries, the ratio of $\frac{L(\vec{Q})}{2S(\vec{Q})}$ can be easily calculated. By comparing the calculated ratio with the Φ averaged experimental values as shown in Fig. 17 (b), $\frac{L(\vec{Q})}{2S(\vec{Q})}$ can be determined. $\frac{L(\vec{Q})}{2S(\vec{Q})}$ as a function of $\vec{Q}$ for three different reflections ($\frac{1}{2}\frac{1}{2}\frac{1}{2}$), ($\frac{3}{2}\frac{3}{2}\frac{3}{2}$), ($\frac{5}{2}\frac{5}{2}\frac{5}{2}$) is shown in Fig. 18(a). The results in Fig. 18 clearly show that a large contribution (17 $\pm$ 3%) to the magnetization from the orbital moment exists in NiO. The increase of $L(\vec{Q})/2S(\vec{Q})$ with the scattering vector reflects the broader spatial extent of the spin density. The orbital contribution enhances the spin-alone magnetic moment to make the total moment eventually larger than 2 μ_B as expected for spin-only magnetic moment (*S* = 1).

The orbital and spin angular momentum can be put in an absolute scale by comparing scattered magnetic intensities with charge intensities and by talking care of extinction correction, if any. The resulting values are shown in Fig. 18(b) as a function of the scattering vector. The extrapolated values at zero scattering vector, $S(0) = 0.95 \pm 0.1$ and $L(0) = 0.32 \pm 0.05$, lead to a value of $2.2 \pm 0.2 \mu_B$ for the staggered magnetization at T = 300 K. This is in close agreement with neutron results giving 1.81 ± 0.2 μ_B [52] and 1.97 μ_B [53].

Summarizing the main results of NiO, (a) the spin and orbital contributing to the total magnetization was measured using nonresonant magnetic scattering. The extrapolation at $\vec{Q}$ = 0 shows that the effective *L*/*S* ratio amounts to 0.34. This is unexpectedly large contribution in transition metal oxides like NiO where *L* is supposed to be largely quenched. Furthermore the orbital angular momentum is parallel to spin, as expected from simple spin-orbit coupling for a $3d^8$ free atom. The results also indicate a contraction of the atomic wave function for Ni in NiO, if the experimental values are compared with the predictions of theory. The atomic wave functions are contracted by 17 % as compared to the free ion. We can conclude that the magnetization density in a simple system like NiO is not yet fully understood and the results of these studies have to be taken into account, when models for the electronic and magnetic structure and properties of transition metal compounds are being made.

5.3 Nonresonant magnetic scattering from Ho Metal:

Nonresonant magnetic scattering from Ho metal illustrates the advantages of magnetic x-ray scattering compared to the neutron scattering and reveals new features in magnetic structure that was unobserved in neutron diffraction experiments [4, 56, 57]. Ho metal has an h.c.p. crystal structure with two layers per chemical unit cell and a large magnetic moment of $10\mu_B$/atom. Below the magnetic ordering temperature, $T_N \approx$ 131K, a pair of satellite reflection appears around each main Bragg reflection parallel to the reciprocal *c* axis [57, 58]. These results are usually taken as evidence for a simple spiral antiferromagnetic structure in which the average

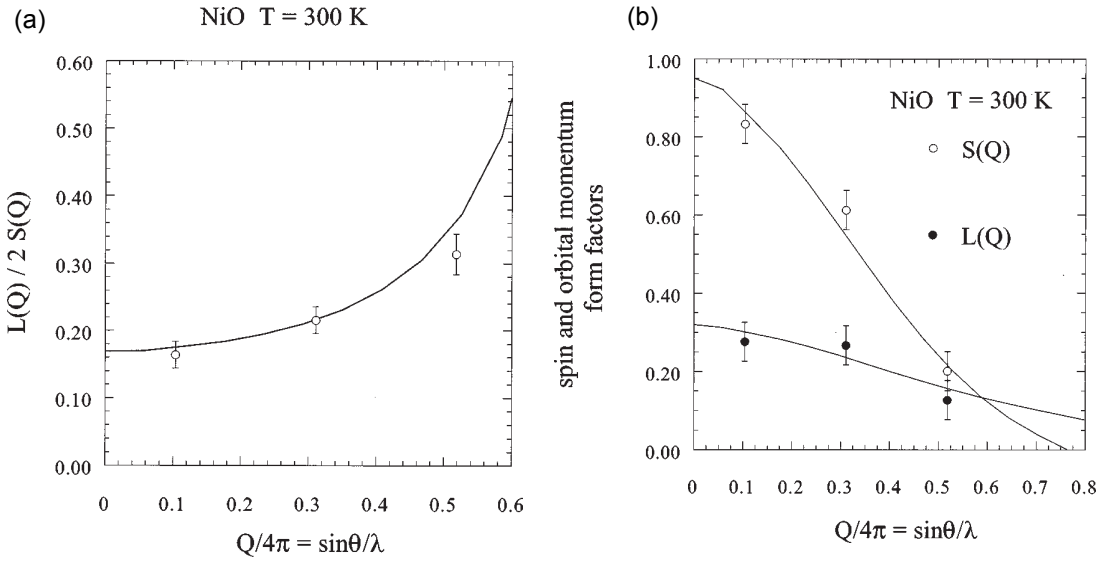


Figure 18: *(a) Measured variation of* $L(\vec{Q})/2S(\vec{Q})$ *as a function of* $\sin\frac{\theta}{\lambda} = \frac{Q}{4\pi}$. *The continuous line is the K dependence estimated by Blume [Ref. [54]] adjusted to fit through the data with a contraction of the wave function by 17%.* $L(\vec{Q})/2S(\vec{Q})$ *extrapolates to 0.17 at* $\vec{Q} = 0$. *(b) Spin form factor and orbital-moment form factor in NiO. The data have been obtained by normalizing magnetic intensities to charge peaks corrected for extinction. The continuous lines are the calculated variations of* $S(\vec{Q})$ *and* $L(\vec{Q})$ *with* $\sin\frac{\theta}{\lambda}$ *from Refs. [54] and [55] with an expansion of the* $\vec{Q}$ *scale by 17%.*

moments are aligned ferromagnetically within the basal planes, but rotate from one plane to the next plane with an average turn angle varying between ≈50°/layer at T_N and ≈30°/layer near T_c. The exact turn angle can be determined by measuring the modulation wave vector at each temperature. Figure 19 shows the schematic of the magnetic structure along with the temperature dependence of the magnetic propagation vector.

Gibbs *et al.* [4] first reported non-resonant x-ray magnetic scattering on Ho. The count rate of the magnetic signal was comparable with neutron diffraction (~25 counts/sec). The background of the NRXMS was one order of magnitude higher than that of neutron. This was more than compensated for by a five fold increase of resolution, which proved to be very important for the success of the experiment. The wave-vector resolution of 10^{-3} Å^{-1} reveals new features in the temperature dependence of the propagation vector and leads to a new model of the magnetic structure. The temperature dependence of the propagation vector measured by x-rays is in very good agreement with the neutrons as is shown in Fig. 19(a). In addition, several lock-in transition was observed as shown in the inset of Fig. 19(a). Below 50 K thermal hysteresis, irreversibility, coexistence of phases with differing wave vector, and apparent lock-in behavior was observed.

Figure 20 (a) shows the diffraction pattern obtained with synchrotron radiation at 17 K when the magnetic satellite is located at a commensurate position $\tau_m = \frac{5}{27}$. A second peak of intensity comparable with the satellite but of greater width (0.0075 vs 0.005 Å^{-1}) is observed at $\tau_c = \frac{2}{9}$. As the temperature is increased, τ_m shifts away from $\frac{5}{27}$, and the second peak apparently broadens and disappears. To establish the origin of the additional scattering at $\frac{2}{9}$, the scan at 17 K was repeated with the polarization analyzer in place. As seen in Fig. 20(a), the second peak is completely eliminated when measured with PG(0 0 6) analyzer crystal in the $\sigma \rightarrow \pi'$

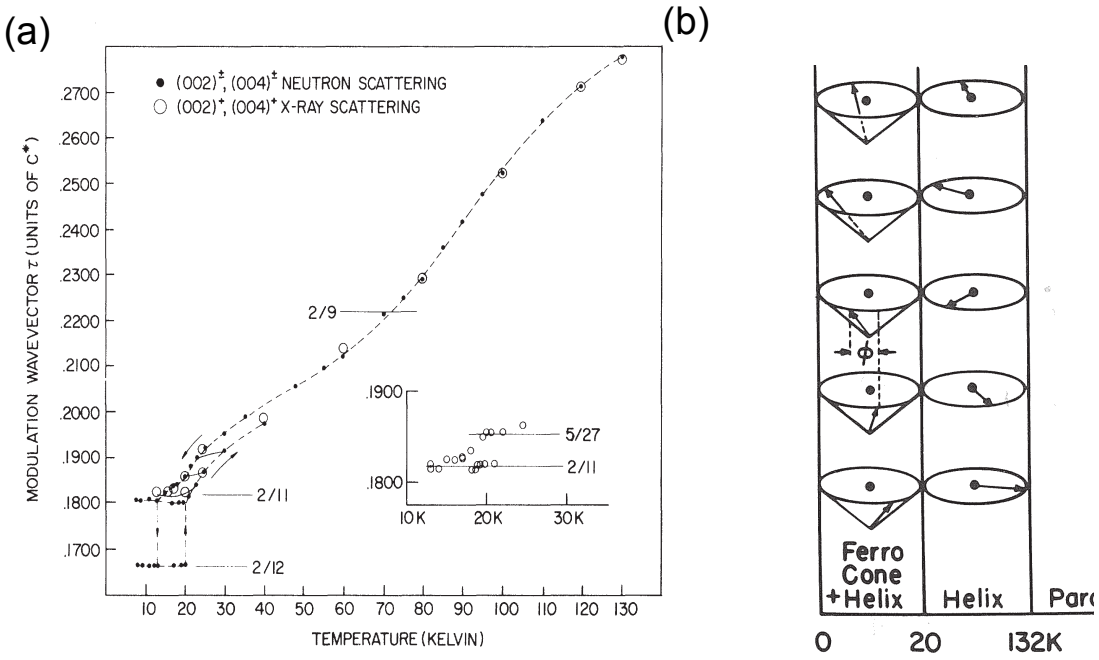


Figure 19: *(a) Temperature dependence of the Ho modulation wave vector τ obtained with both synchrotron x-ray (open circles) and neutron (filled circles) scattering. The wave vectors obtained by neutron scattering in the coexistence region below 20 K are the result of fits to the first harmonic. Th fine lines across the hysteresis loop indicate the results of cycles of the temperature below 50 K. Inset: Plot of the wave vectors obtained from several cycles of the temperature between 13 and 24.5 K obtained with x-ray scattering. [4] (b) Schematic of the magnetic structure of Ho determined using neutron diffraction as a function of temperature. [7].*

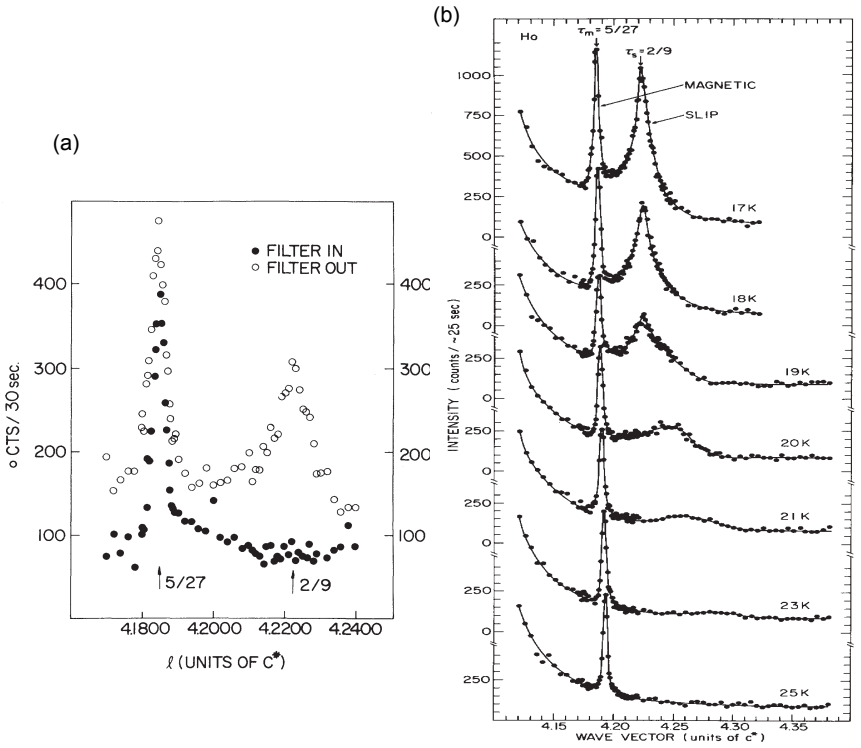


Figure 20: *(a) Open circles: Scan of the Ho(004) magnetic ($\tau = \frac{5}{27}$) and charge ($\tau = \frac{2}{9}$) satellites taken at 17 K. Filled circles: The same scan, but with the polarization analyzer in place.(b) Temperature dependence of the magnetic and charge satellites measured during cooling.[59]*

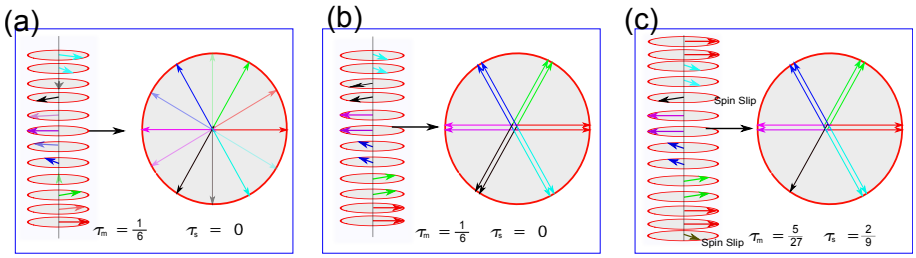


Figure 21: *(a) The magnetic structure proposed by neutron scattering at low temperature where moments in the consecutive layers rotates by 30°. (b) The low temperature magnetic structure as proposed by Gibbs et al. [4] where a pair of doublet rotates by 60° giving the same propagation vector as (a). (c) The spin-slip model for the magnetic structure. In(a)-(c), the right sides depicts projection of the magnetic moments in the hexagonal a-b plane.*

scattering channel, showing unambiguously that (1) the scattering at τ_m=$\frac{5}{27}$, is magnetic in origin, and (2) the peak at τ_c=$\frac{2}{9}$, originates from the charge scattering. For a pure spiral with the propagation vector perpendicular to the moment direction, magnetoelastic coupling is zero and thus, no periodic modulation of the lattice is expected [7]. Therefore, the existence of the charge scattering at $\tau_c = \frac{2}{9}$ is completely unexpected from the previous model of the magnetic structure.

The existence of the commensurate magnetic satellite and associated charge scattering can be explained based on the spin slip model. Spin slip model of the magnetic structure is based on the low temperature τ_m= $\frac{1}{6}$ structure which is schematically shown in the Fig. 21. In this structure the pairs of magnetic moments or doublets are oriented along one of the easy directions of hexagonal lattice. The adjacent pairs are rotated by 60° so that 12 atomic layers or 6 chemical unit cells are needed to complete a 360°turn. Thus the magnetic period τ_m= $\frac{1}{6}$. Shorter period modulations occur due to periodic spin slips where single atomic layer (singlet) makes a 60° turn as shown in Fig. 21(c). In principle, the spin-slip model allows a large number of possible sequences of doublet and singlet, producing modulation wavevector in the range $\frac{1}{3} \geq \tau_m \geq \frac{1}{6}$. The change in magnetoelastic coupling at spin-slip sites causes a small distortion of the chemical lattice, with a period equal to the spin slip period. The spin-slip model describes not only the lock-in behavior of the magnetic modulation wave vector, but also predicts also the wave vector for the charge scattering. In Ho it was proposed that the magnetic structure consists of two basic blocks: the spin-slip block consisting of single atomic layer and a no-spin-slip block consisting of a doublet. A magnetic unit cell then can be described as a sequence of *N* no-slip blocks and *S* spin-slip blocks. For the periodic sequence of *N* ans *S* blocks it can be shown in the case of Ho metal that: [7]

$$\tau_m = \frac{m+1}{6m+3} \tag{41}$$

$$\tau_c = \frac{2}{2m+1} \tag{42}$$

where $m = \frac{N}{S}$. Some of the allowed wave-vectors for different values of *m* are given in Table 4. One can see that there is a remarkable agreement between the allowed wave-vectors in the spin-slip model and the observed experimental data shown in Fig. 19(a). Although we have assumed a periodic structure for the spin-slip array, it is clear from the data of Fig. 20 that the peak at $\frac{2}{9}$ is not instrumentally narrow. This width is direct evidence of the lack of long range periodicity in the spin-slip distribution.

Table 4: *Some of the commensurate magnetic (τ_m) and lattice modulation wave vectors predicted by the spin-slip model of Ho [7]. The rotation of magnetic moment between no-spin-slip block and spin-slip block is 60°.*

m	0	1	2	3	4	5	6	∞
τ_m	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{5}$	$\frac{4}{21}$	$\frac{5}{27}$	$\frac{2}{11}$	$\frac{7}{39}$	$\frac{1}{6}$
τ_s	2	$\frac{2}{3}$	$\frac{2}{5}$	$\frac{2}{7}$	$\frac{2}{9}$	$\frac{2}{11}$	$\frac{2}{13}$	0

So far we have discussed nonresonant magnetic diffraction at low to medium energies (4-20 keV). The nonresonant magnetic diffraction can be performed also at very high energies which has several advantages. First of all, instead of specialized Be windows, less expensive Aluminum windows can be used due to less absorption at high energies. Secondly, due to the high

penetration of the high energy x-rays, bulk sample can be probed in transmission geometry and one do not have to worry about sample surface preparation. For 3*d* transition metals, the absorption length increases from some μm at 8 keV to several mm at 80 keV. Due to the enhancement of penetration depth, the effective scattering volume increases and hence, the scattering intensity increases by a factor of 3-4 orders of magnitude. Corrections for absorption, extinction, beam foot print etc. are simple and therefore, by normalizing the intensity of the magnetic reflections to the intensity of the charge reflections, absolute values for the spin moment can be determined [60]. For high energies, scattering angles are very small for the low index reflections. Therefore, neglecting terms in $\sin^2\theta$ one can write the nonresonant magnetic cross-section as follows:

$$\begin{aligned} \langle M_m \rangle &= \begin{pmatrix} A_{\sigma\to\sigma'} & A_{\sigma\to\pi'} \\ A_{\pi\to\sigma'} & A_{\pi\to\pi'} \end{pmatrix} \\ &= \begin{pmatrix} (\sin 2\theta)S_2 & -0 \\ 0 & (\sin 2\theta)S_2 \end{pmatrix} \end{aligned} \tag{43}$$

We can see from the matrix Eq. 43 that the magnetic scattering at high energies is only sensitive to spin component perpendicular to the scattering plane and the scattering does not change the polarization state. Nonresonant magnetic scattering at high energies has been successfully employed to study the magnetism of MnF_2 by Strempfer *et al.* [61]. With 80 keV x-rays a high peak count rates 13000 counts/sec with a peak to background ratio of 230:1 and a very high reciprocal space resolution can be obtained for the magnetic (3 0 0) reflection as shown in Fig. 22(a). Figure 22 (b) shows a measurement of the temperature dependence of the sublattice magnetization. In the critical region close to the Neél temperature T_N, the reduced sublattice magnetization $m = M(T)/M(T=0)$ follows very accurately a power law behavior:

$$m(\tau) = D\tau^{\beta} \tag{44}$$

as a function of the reduced temperature $\tau = (T_N - T)/T_N$. The value of the critical exponent of β = 0.333(3) corresponds well to the predictions of the Ising model. As discussed previously, the problem of multiple charge scattering is severe at high energies and unfortunately, it was the case for MnF_2 due to the identical chemical and magnetic unit cell and specific crystal symmetry. Multiple charge scattering was minimized by judiciously selecting the azimuth angle as discussed in section 5.1.1.

5.4 Magnetic scattering from ferromagnet:

We have discussed applications of resonant and nonresonant magnetic scattering in the case of antiferromagnets where magnetic signal is well separated form charge scattering. For a ferromagnetic samples or in an atiferromagnets with $\vec{Q}$ = 0, where charge and magnetic scattering coincide, it seems impossible to measure the magnetic signal. However, it has been shown that for a ferromagnet it is possible to measure magnetic signal by measuring (a) flipping ratio (asymmetry ratio) in a magnetic field or (b) by using the the fact that the charge scattering from a sample is suppressed by a factor of $\cos^2 2\theta_{sample}$ in addition to $\cos^2 2\theta_{analyzer}$ in $\pi \to \sigma'$ scattering geometry.

(a) The total scattering cross-section generated by the magnetic and charge scattering length is:[7]

$$\frac{d\sigma}{d\omega} = |f|^2 = |f_m + f_c|^2 \tag{45}$$

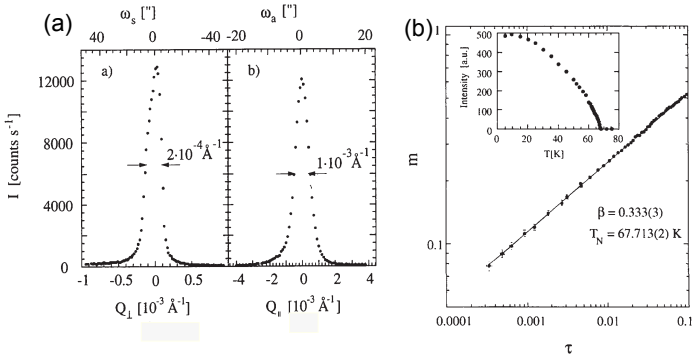


Figure 22: *(a) Representative scans of the magnetic (3 0 0) reflection at T = 5 K. (a) and (c) show transversal scans in the three and two crystal modes, respectively.(b)The main graph shows the reduced magnetization* m=M(T)/M(0) *as a function of the reduced temperture* $\tau = (T_N - T)/T_N$ *on a double logarithmic scale for the magnetic (3 0 0) reflection. The points are taken from the measurement, the solid line represents a fit using equation 44. The inset shows the temperature dependence of the intensity of the magnetic (3 0 0) reflection from 5 to 80 K with* T_N *= 67.713 K. Figure adapted from Ref. [61].*

If f_m and f_c are exactly in phase, then the Eq. 45 can be written as:

$$\frac{d\sigma}{d\omega} = f_c^2 + f_m^2 + 2f_c f_m \tag{46}$$

Eq.46 has three terms: the first term is the pure charge scattering which is the most dominant term. The second term is the pure magnetic term and the third term is due to the interference between charge and magnetic signals. The pure magnetic term is 10^{-6} times weaker than the pure charge term and therefore, undetectable in a magnetic scattering experiment. The third term being linear in f_m, it is much stronger than the pure magnetic term. Moreover, being linear in f_m, the sign of the third term can be changed by changing the magnetization direction by an external magnetic field. Thus, the difference in the cross section for two different orientation of the magnetic field can be detected.

Unfortunately, in a centrosymmetric crystal structure, the f_m and f_c are exactly out of phase for the plane polarized incident light, and therefore, the interference term vanishes. The interference term for the centrosymmetric case only exists if the incident light is elliptically polarized. In a non-centrosymmetric crystal structure the scattering phases depend on the precise distribution of the charges and magnetic moment in the unit cell and depending on the phase difference an interference term is permitted. One clever way to introduce a phase difference is to tune the x-ray energy to the absorption edge of one of the element in the crystal. Resonant scattering introduces a strong energy dependent phase shift and hence an interference term is allowed. Charge magnetic interference was first seen in ferromagnetic Ni single crystal at the Ni *K* edge. By reversing the direction of external magnetic field applied perpendicular to the horizontal scattering plane an asymmetry ratio can be calculated:

$$R_a = \frac{I\uparrow - I\downarrow}{I\uparrow + I\downarrow} \tag{47}$$

Where $I\uparrow$ and $I\downarrow$ represent scattering intensities for two opposite magnetization direction. The value of asymmetry ratio R_a of 10^{-3} and 0.1 was reported for Ni [62]and CoPt alloy [63],

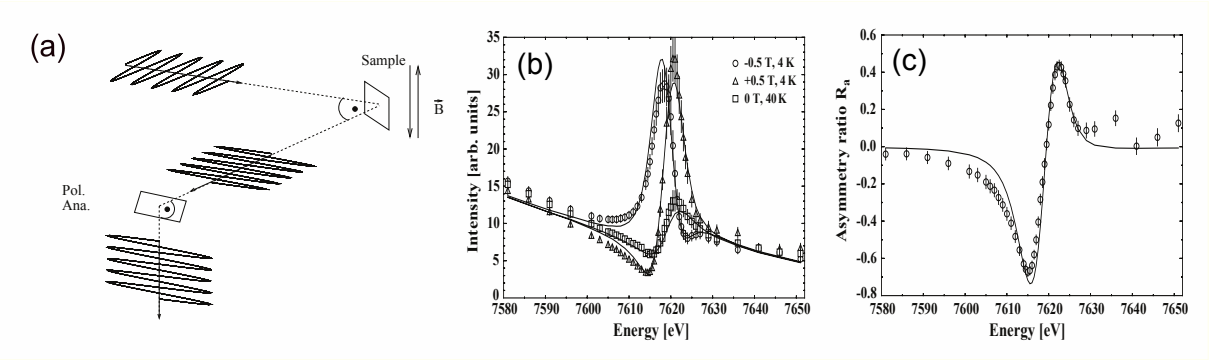


Figure 23: *(a) Experimental set-up for the measurement of XRMS from ferromagnetic samples. The primary beam is polarized horizontally, the magnetic field is applied perpendicular to the horizontal scattering plane. The Bragg peak of the sample and the analyzer were chosen to scatter close to 90 degrees to minimize the charge scattering. (b) Absorption-corrected energy dependence of the (1 1 5) reflection measured at the Eu* L_{II} *edge. Two measurements with magnetic fields of +0.5 T and −0.5 T were performed at 4 K. Charge and magnetic intensities contribute to the scattering amplitude. (c) Asymmetry ratio* R_a *of the (1 1 5) reflection at 4 K at the Eu* L_{II} *edge. Figures adapted from Ref. [64].*

respectively. The asymmetry ratio was also measured at the Eu L_{II} edge for the EuS single crystal and the ratio as large as R_a of 0.67 was obtained [64].
Fig. 23 (a) shows the experimental geometry for measuring the flipping ratio for the EuS single crystal. The (1 1 5) reflation was measured in the $\pi \rightarrow \pi'$ scattering geometry to suppress charge signal relative to the magnetic scattering signal and to access the magnetic moment component perpendicular to the scattering plane. In the $\pi \rightarrow \pi'$ scattering geometry, resonant scattering amplitude is sensitive to the magnetic moment perpendicular to the scattering plane i.e. in the direction of the applied magnetic field (m_z). Therefore, we can write:

$$I = I_{mag} \left\langle m_z^2 \right\rangle + I_{int} \left\langle m_z \right\rangle + I_{charge}$$

where I_{mag} is the pure magnetic term, I_{charge} is the charge term and I_{int} is the interference term and $\langle m_z \rangle$ is the expectation value of the magnetic moment along the field direction. Figure 23(b) shows the resonance spectra at the Eu L_{II} edge below and above the ferromagnetic transition temperature of 17 K. The resonance spectra below T_c was measured for two opposite field directions. The asymmetry ratio can be calculated using the formula 47 and is shown in Fig. 23 (c). Clearly the asymmetry ratio is quite large compared to Ni and CoPt. Important spectroscopic information such as the exchange splitting of the 5*d* conduction band can be obtained by fitting the scattered intensity on the energy and the applied magnetic field [64]. The exchange splitting was found to be $\epsilon = 0.27(1)$eV.
Ferrimagnetism of Gd moments in $GdNi_2Ge_2$ [65] as well as antiferromagnetism with Q = 0 magnetic structure in SmFeAsO [34] were studied in zero magnetic field using XRMS at the Gd L_{II} and Sm L_{II} edges, respectively. In both cases, the authors employed $\pi \rightarrow \sigma'$ scattering geometries where the charge signal can be reduced by $\cos^2 2\theta_{analyzer} \times \cos^2 2\theta_{sample}$ compared to the reduction by $\cos^2 2\theta_{analyzer}$ in the $\sigma \rightarrow \pi'$ scattering geometry. In the $\pi \rightarrow \sigma'$ scattering geometry with scattering angle of the sample ($2\theta_{sample}$) close to 90° the charge scattering is reduced by a factor of 10^{-6}. Therefore, the charge and magnetic scattering intensities become comparable and measurement of magnetic signal becomes feasible.

In the above examples, resonant magnetic term interfere with the charge scattering and is known as resonant magnetic-charge interference. Similarly one can utilize the interference term between nonresonant magnetic and resonant charge scattering by tuning the x-ray energy to the absorption edge of a nonmagnetic element in the crystal. This is known as magnetic-resonant charge interference scattering. One can also go far above the absorption edge associated with a weak resonance enhancement (such as *K*-edges of transition metal) so that the resonance enhancement is small and still there is a phase shift which leads to nonzero interference term. This approach was utilized in the first magnetic scattering experiments on ferromagnets [63].

6 Summary

The above examples clearly demonstrates that magnetic x-ray scattering becomes a microscopic probe of magnetism. Like any other experimental techniques magnetic x-ray scattering has it's own weaknesses and strengths. For example, determination of a completely unknown magnetic structure is very difficult since magnetic signal from powder sample is very weak compared to the charge signal. Therefore, magnetic structure determination from powder samples will remain a typical task for neutron scattering where nuclear and magnetic signal have comparable intensities. It is also very difficult to determine the absolute value of the ordered magnetic moment using nonresonant x-ray scattering. Nevertheless, using nonresonant x-ray magnetic scattering, it is possible to separate spin and orbital angular momentum which is not possible using neutron diffraction. Furthermore, in contrast to neutron scattering, resonant scattering is not only element specific but also band selective as shown in the case of $TbMn_2O_5$. Due to the intrinsic collimation of the synchrotron x-rays, the Q-space resolution is much better for a synchrotron experiment. This is often helpful to study incommensurate magnetic structure and detect lock-in transitions as in the case of Ho metal. Magnetic x-ray scattering provides an alternative to study magnetism of compounds which contains strongly neutron absorbing materials such as Gd, Sm,Eu etc. The magnetic form factors for the XRMS, NRXMS and neutron scattering are quite different. In XRMS, the spatial extension of the core levels is relevant and therefore virtually no decrease of the scattering amplitude as a function of momentum transfer is observed. In non-resonant x-ray scattering, the form factors of spin and angular momentum can be determined separately, while neutrons are sensitive to a combination of both.
In summary, both magnetic x-ray and neutron scattering are complementary probe of magnetism and one has to select both methods or any one of them depending on the specific problem.

References

[1] P. M. Platzman and N. Tzoar, Phys. Rev. B **2**, 3556 (1970).

[2] F. D. Bergevin and M. Brunel, Phys. Lett. A **39**, 141 (1972).

[3] F. de Bergevin and M. Brunel, Acta Crystallogr. **A37**, 314 (1981).

[4] D. Gibbs, D. E. Moncton, K. L. D'Amico, J. Bohr, and B. H. Grier, Phys. Rev. Lett. **55**, 234 (1985).

[5] D. Gibbs, D. R. Harshman, E. D. Isaacs, D. B. McWhan, D. Mills, and C. Vettier, Phys. Rev. Lett. **61**, 1241 (1991).

[6] E. Balcar and S. Lovesey, *Theory of magnetic neutron and photon scattering* (Oxford University Press, 1989).

[7] S. Lovesey and S.P.Collins, *X-ray scattering and absorption by magnetic materials* (Oxford Science Publications, 1996).

[8] T. Brückel, lecture notes on "Magnetic X-Ray Scattering".

[9] M. Blume, J. Appl. Phys. **57**, 3615 (1985).

[10] M. Blume and D. Gibbs, Phys. Rev. B **37**, 1779 (1988).

[11] J. P. Hill and D. F. McMorrow, Acta Crystallogr. **A52**, 236 (1996).

[12] J. P. Hannon, G. T. Trammell, M. Blume, and D. Gibbs, Phys. Rev. Lett. **61**, 1245 (1988).

[13] J. W. Kim, Y. Lee, D. Wermeille, B. Sieve, L. Tan, S. L. Bud'ko, S. Law, P. C. Canfield, B. N. Harmon, and A. I. Goldman, Phys. Rev. B **72**, 064403 (2005).

[14] C. Detlefs, A. H. M. Z. Islam, A. I. Goldman, C. Stassis, P. C. Canfield, J. P. Hill, and D. Gibbs, Phys. Rev. B **55**, R680 (1997).

[15] C. Detlefs, Ph.D. thesis, Iowa State University, Ames, Iowa, USA (1997).

[16] M. D. Hamrick, M. A. Thesis, Rice University (1990).

[17] J. P. Hill and D. F. McMorrow, Acta Cryst. **A52**, 236 (1996).

[18] A. I. Goldman, K. Mohanty, G. Shirane, P. M. Horn, R. L. Greene, C. J. Peters, T. R. Thurston, and R. J. Birgeneau, Phys. Rev. B **36**, 5609 (1987).

[19] T. Brückel, M. Lippert, T. Köhler, J. R. Schneider, W. Prandl, V. Rilling, and M. Schilling, Acta Crystallogr. **A52**, 427 (1996).

[20] J. P. Hill, Q. Feng, R. J. Birgeneau, and T. R. Thurston, Z. Phys. B **92**, 285 (1993).

[21] D. Mannix, P. de Camargo, C. Giles, A. de Oliveira, F. Yokaichiya, and C. Vettier, Eur.. Phys. J. B **20**, 19 (2001).

[22] M. K. Sanyal, D. Gibbs, J. Bohr, and M. Wulff, Phys. Rev. B **49**, 1079 (1994).

[23] J. Als-Nielsen and D. McMorrow, *Elements of Modern X-ray Physics* (John Wiley & Sons, New York, 2001).

[24] J. P. Hill, *Magnetic x-ray scattering* (John Wiley & Sons, New York, 2002).

[25] C. Giles, C. Vettier, F. de Bergevin, C. Malgrange, G. Grubel, and F. Grossi, Rev. Sci. Instrum. **66**, 1518 (1995).

[26] L. Bouchenoire, S. D. Brown, P. Thompson, J. A. Duffy, J. W. Taylor, and M. J. Cooper, J. Synchrotron Radiat. **10** (2003).

[27] C. M. C. Detlefs, M. Sanchez del Rio, arXiv:1106.4446 (2011).

[28] V. K. Pecharsky and K. A. Gschneidner, Jr., Phys. Rev. Lett. **78**, 4494 (1997).

[29] V. K. Pecharsky and J. K. A. Gschneidner, Appl. Phys. Lett. **70**, 3299 (1997).

[30] L. Morellon, J. Stankiewicz, B. García-Landa, P. A. Algarabel, and M. R. Ibarra, Appl. Phys. Lett. **73**, 3462 (1998).

[31] L. Morellon, J. Blasco, P. A. Algarabel, and M. R. Ibarra, Phys. Rev. B **62**, 1022 (2000).

[32] L. Tan, A. Kreyssig, J. W. Kim, A. I. Goldman, R. J. McQueeney, D. Wermeille, B. Sieve, T. A. Lograsso, D. L. Schlagel, S. L. Budko, et al., Phys. Rev. B **71**, 214408 (2005).

[33] S.-L. Chang, *Multiple Diffraction of X-Rays in Crystals* (Springer-Verlag, 1984).

[34] S. Nandi, Y. Su, Y. Xiao, S. Price, X. F. Wang, X. H. Chen, J. Herrero-Martin, C. Mazzoli, H. C. Walker, L. Paolasini, et al., Phys. Rev. B **84**, 054419 (2011).

[35] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, J. Am. Chem. Soc. **130**, 3296 (2008).

[36] M. Rotter, M. Tegel, and D. Johrendt, Phys. Rev. Lett. **101**, 107006 (2008).

[37] A. Martinelli, A. Palenzona, C. Ferdeghini, M. Putti, and H. Emerich, J. Alloy. Compd. **477**, L21 (2009).

[38] A. S. Wills, Physica B **276-278**, 680 (2000).

[39] N. Hur, S. Park, P. Sharma, J. Ahn, S. Guha, and S. Cheong, Nature **429**, 392 (2004).

[40] L. C. Chapon, G. R. Blake, M. J. Gutmann, S. Park, N. Hur, P. G. Radaelli, and S.-W. Cheong, Phys. Rev. Lett. **93**, 177402 (2004).

[41] G. R. Blake, L. C. Chapon, P. G. Radaelli, S. Park, N. Hur, S.-W. Cheong, and J. Rodriguez-Carvajal, Phys. Rev. B **71** (2005).

[42] R. D. Johnson, S. R. Bland, C. Mazzoli, T. A. W. Beale, C.-H. Du, C. Detlefs, S. B. Wilkins, and P. D. Hatton, Phys. Rev. B **78**, 104407 (2008).

[43] C. Vecchini, L. C. Chapon, P. J. Brown, T. Chatterji, S. Park, S.-W. Cheong, and P. G. Radaelli, Phys. Rev. B **77**, 134434 (2008).

[44] V. Fernandez, C. Vettier, F. de Bergevin, C. Giles, and W. Neubeck, Phys. Rev. B **57**, 7870 (1998).

[45] L. D. Collins, S. P. and G. H. Guo, J. Phys. Condens. Matter **5**, L637 (1993).

[46] C. S. Laundry, D. and A. J. Rollason, J. Phys. Condens. Matter. **3**, 369 (1991).

[47] B. Brandow, Adv. Phys. **26**, 651 (1977).

[48] K. Terakura, A. R. Williams, T. Oguchi, and J. Kübler, Phys. Rev. Lett. **52**, 1830 (1984).

[49] K. Terakura, T. Oguchi, A. R. Williams, and J. Kübler, Phys. Rev. B **30**, 4734 (1984).

[50] C. G. Shull, W. A. Strauser, and E. O. Wollan, Phys. Rev. **83**, 333 (1951).

[51] J. Baruchel, M. Schlenker, K. Kurosawa, and S. Saito, Philos. Mag. B **43**, 853 (1981).

[52] H. A. Alperin, Phys. Rev. Lett. **6**, 55 (1961).

[53] B. E. F. Fender, A. J. Jacobson, and F. A. Wedgwood, J. Chem. Phys. **48**, 990 (1968).

[54] M. Blume, Phys. Rev. **124**, 96 (1961).

[55] R. E. Watson and A. J. Freeman, Acta Crystallogr. **14**, 27 (1961).

[56] W. C. Koehler, J. W. Cable, H. R. Child, M. K. Wilkinson, and E. O. Wollan, Phys. Rev. **158**, 450 (1967).

[57] W. C. Koehler, J. W. Cable, M. K. Wilkinson, and E. O. Wollan, Phys. Rev. **151**, 414 (1966).

[58] M. J. Pechan and C. Stassis, J. Appl. Phys. **55**, 1900 (1984).

[59] J. Bohr, D. Gibbs, D. Moncton, and K. D'Amico, Physica A **140**, 349 (1986).

[60] J. Strempfer, U. Rütt, and W. Jauch, Phys. Rev. Lett. **86**, 3152 (2001).

[61] J. Strempfer, T. Brückel, U. Rütt, J. R. Schneider, K.-D. Liss, and T. Tschentscher, Acta Crystallogr. **A52**, 438 (1996).

[62] K. Namikawa, M. Ando, T. Nakajima, and H. Kawata, J. Phys. Soc. Jpn. **54**, 4099 (1985).

[63] F. de Bergevin, M. Brunel, R. M. Galera, C. Vettier, E. Elkaim, M. Bessiere, and S. Lefebvre, Phys. Rev. B **46**, 10772 (1992).

[64] D. Hupfeld, O. H. Seeck, J. Voigt, J. Bos, K. Fischer, and T. Brückel, Europhys. Lett. **59**, 284 (2002).

[65] J. W. Kim, A. Kreyssig, P. Ryan, E. Mun, P. C. Canfield, and A. I. Goldman, Appl. Phys. Lett. **90**, 202501 (2007).

C 8 The European X-ray Free-Electron Laser Project [1]

M. Altarelli
European XFEL GmbH
22607 Hamburg

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Synchrotron radiation sources have revolutionized UV and x-ray experiments in many fields of science. The driving force behind the development of light sources is the optimization of their brilliance (or spectral brightness), which is the figure of merit for many experiments. Brilliance is defined as a function of frequency given by the number of photons emitted by the source in unit time in a unit solid angle, per unit surface of the source, and in a unit bandwidth of frequencies around the given one.

In the most modern synchrotron sources (the so-called "third-generation light sources", such as the ESRF, Elettra, Diamond, Swiss Light Source, etc.) the average brilliance of undulator radiation reaches values up to $10^{19} - 10^{20} photons/s/mrad^2/mm^2/0.1\%BW$ [1]. Taking into account the pulsed nature of the sources, i.e. the filling patterns and revolution times of storage rings, this corresponds to peak brilliance values of $10^{24} photons/s/mrad^2/mm^2/0.1\%BW$. In order to achieve such values, two ingredients are essential. The first ingredient is the extensive use of undulators as radiation sources. An undulator is a set of two arrays of magnets, located in a straight section of the ring above and below the vacuum chamber where electrons run, and subjecting the electrons to a vertical magnetic field varying with position in a sinusoidal manner. The corresponding Lorentz force on the electrons results in an oscillating trajectory, with many bending points from which emission of synchrotron radiation occurs. In undulators, the broadband radiated power of bending magnet radiation, due to the interference of the different emission points along the trajectory, is concentrated in a spectrum of narrow lines, centered about the wavelengths:

$$n\lambda = (\lambda_u/2\gamma^2)(1 + K^2/2) \tag{1}$$

Here $n = 1, 2, 3 \ldots$ is the order of the harmonic, λ_u is the period of the undulator magnetic structure, γ is the electron energy, expressed in units of the electron rest energy, and K is the undulator parameter, a number of order 1 given by $K = \gamma\theta$, where θ is the maximum angular deviation of the electrons from their unperturbed trajectory, induced by the undulator magnetic field. It can be shown that Eq. (1) identifies the wavelength of the fundamental harmonic λ as the distance by which one electron lags behind the emitted photons after traveling over the distance λ_u from the emission point.

The second ingredient is the reduction in the phase-space volume of the circulating electrons in the two transverse directions (the horizontal and vertical directions perpendicular to the average orbit). These quantities are called horizontal (vertical) emittances and are roughly speaking a measure of the horizontal (vertical) size of the electron bunch times the angular divergences of the corresponding velocity vectors projections. Progress in accelerator physics has allowed reduction of the horizontal emittance to values of order $1\,nm\,rad$, as presently achieved by the $6\,GeV$ Petra III ring at DESY [2]. It is intuitive that the properties of small dimension and high collimation of the electron beam translate into corresponding attributes of the radiated photons, and therefore in higher brilliance. A substantial further reduction of emittance values towards the fundamental limits is presently the subject of extensive research on the so called "ultimate" storage ring source [3]. Rings under construction, such as the $3\,GeV$ NSLS II in Brookhaven [4] and Max IV in Lund, Sweden [5], aim at emittances of a few $0.1\,nm$, and machines on the drawing board, such as the $4.5\,GeV$ PEP-X project at SLAC in Stanford [6] aim at the $0.1\,nm$ and below range. In general, the achievement of very low emittances involves the use of long rings (both Petra III and PEP exceed $2\,km$ in circumference), special devices such as damping

wigglers, and daring innovations in lattice design, with compromises on maximum current, lifetime and ease of injection.
The light source becomes diffraction limited for radiation at wavelength λ when the emittance is reduced towards and beyond $\lambda/4\pi$. This means that there is a high degree of transverse coherence in the radiation at that wavelength.
Another fundamental limitation of storage rings concerns the bunch length, i.e. the duration of the light pulses. Typically, pulse duration in storage rings is limited to some $30\,ps$; if the bunch length is compressed below such a length, the bunch will anyway return to the original length after a certain number of turns, because of the intense emission processes in the insertion devices and in the bending magnets; the equilibrium length results from a balance between quantum excitations and radiation damping [3]. Substantially shorter pulses can only be achieved at the expense of dramatic reductions of the radiated intensity. This poses a limitation to the time scales which can be explored by time-resolved experiments with synchrotron sources: at full power they are limited to the $\simeq 50\,ps$ time scale; access to the scale of atomic motions and rearrangements (typically, sub-ps), is only possible by techniques such as "bunch slicing", which produce pulses of $100\,fs$, but with intensities limited to $\simeq 10^3$ photons per pulse, and a few kHz pulse repetition rate [7]. On the other hand, there is a high demand for ultrafast experiments capable to explore atomic motions and configuration changes on a sub-ps time scale. The development of fs lasers in the infrared, the visible and near UV has shown a variety of interesting phenomena essential for the understanding of chemical reactions, phase transitions, etc.; only shorter wavelengths, however, can resolve smaller and smaller distances, and ultimately only x-rays can provide us with atomic position information.
In the following sections we shall review progress in the realization of x-ray FEL (free-electron laser) sources, based on linear accelerators, which allow generation of transversely coherent ultrashort (typically $10 - 100\,fs$) pulses, with a spectacular increase of some nine orders of magnitude in peak brilliance with respect to third-generation synchrotron sources (see Fig. 1). The linear accelerator allows indeed to obtain very low emittances, and, in addition, is a single-pass machine, in which the electron bunches run only once through the undulator, keeping the original bunch length.

2 The SASE process and single-pass Free-Electron Lasers

In the undulators of a synchrotron source, electrons are forced to follow a zigzag trajectory by the device magnetic field. There is a definite phase relationship between the radiation emitted by the same electron at different points of the trajectory, and, since the fields overlap (the angle θ of maximum deviation, entering the undulator parameter K of Eq. (1), is of order $1/\gamma$, i.e. of the aperture of the radiation cone) there is an interference, which is constructive only for the wavelengths described by Eq. (1). Notice, however, that under such circumstances all interference between the fields radiated by different electrons is averaged out, as no definite phase relationship occurs between them. The reason is that electrons are randomly distributed inside the bunch, with no correlation between positions of different electrons. In order to have such interference, electrons should be spatially ordered; considering for simplicity two electrons, if the longitudinal coordinate (projection on the undulator axis of the position) of the second lags behind that of the first by an integer number of wavelengths, the corresponding radiation fields will superpose in phase after the electrons have run through an integer number of undulator periods: remember that the light radiated by the second will "catch up" with the first, getting

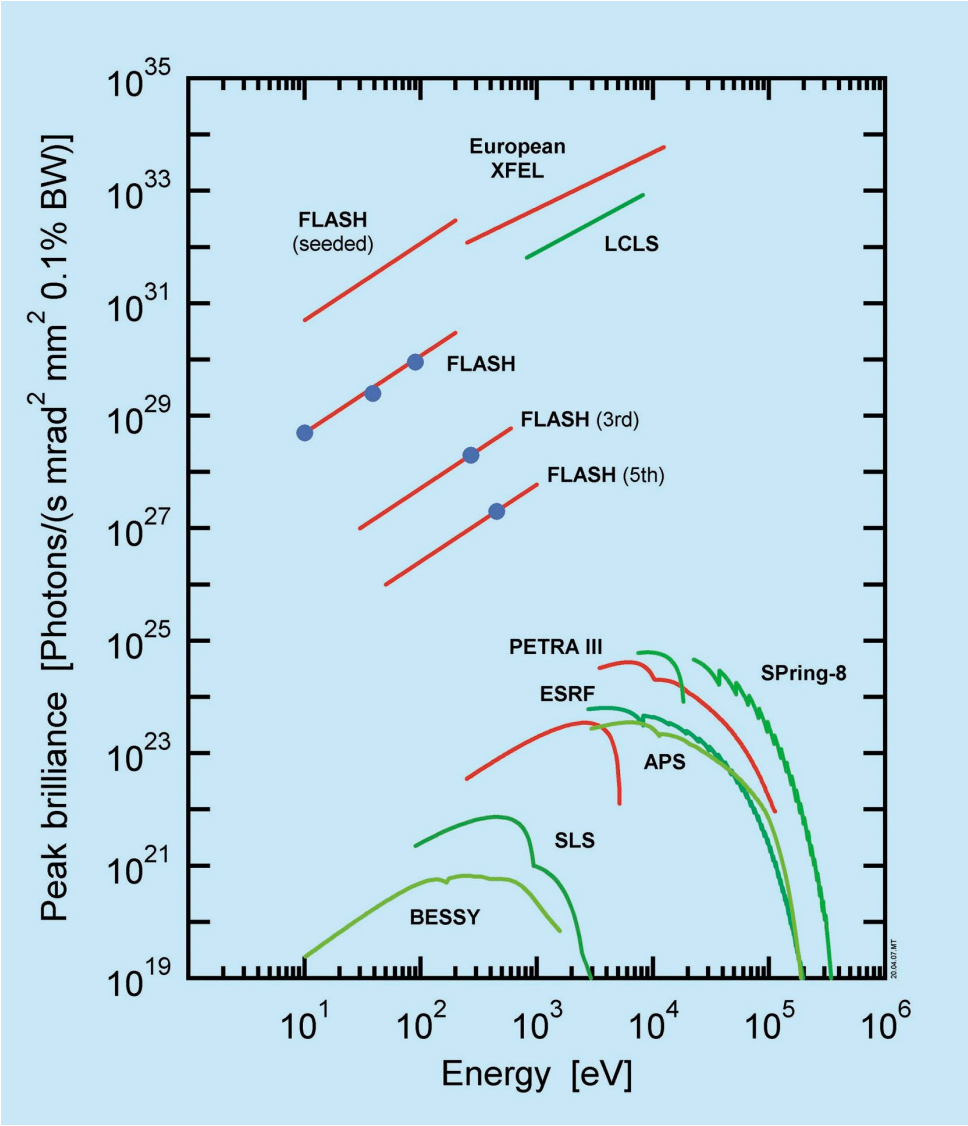


Fig. 1: *Peak brilliance as a function of photon energy of FLASH, the LCLS in Stanford and the future European XFEL in Hamburg (an upgraded version of FLASH based on "seeding", see Section 5, is also shown). For comparison, some 3rd generation synchrotron radiation facilities are shown. Dots denote measured values at FLASH. The third and fifth harmonics of the FLASH undulator, on which lasing was observed, but not saturation, are also shown.*

closer to it at the rate of one wavelength λ per undulator period λ_u. The intensity radiated from the two electrons will be four times larger than that of one single electron. From these simple considerations one can understand how coherence effects between different electrons can arise when the density in the bunch (integrated over the transverse directions) has a Fourier component at the wavelength of the radiation, i.e. when this density shows a modulation at the radiation wavelength. The intensity of the radiation in such cases becomes proportional to the square of the number of electrons involved in the modulation.

For short wavelengths, in the nm range or below, controlling the electron density on that scale may appear extremely difficult. However, in a certain sense, the radiation does it for us. This microbunching phenomenon occurs because the electric field of the radiation has a small component parallel (or antiparallel) to the electron velocity (see Fig. 1), which tends to accelerate some electrons and decelerate those which are positioned one half radiation wavelength ahead or behind, leading to bunching on the radiation wavelength scale. Whenever shot-noise fluctuations in the electron bunch introduce a Fourier component of the appropriate wavelength in the electron density, the coherence effect between electrons described above increases the radiated intensity; it turns out that, for a sufficiently low-emittance and high peak-current electron beam, in a sufficiently long undulator, the stronger radiation field, via the microbunching process, reinforces the density fluctuation, and so on, in a runaway process that leads to exponential amplification of the radiated intensity. The amplification proceeds until saturation, which occurs when the intense radiation and subsequent recoil effects lead to a degradation of the electron beam quality that prevents further amplification. This single-pass process, known as Self-Amplified Spontaneous Emission (SASE) was theoretically identified many years ago, long before electron beams of sufficient density and quality were technologically feasible [8], [9], [10].

The first experimental demonstration was in the visible range, at the LEUTL facility at Argonne National Laboratory [11], and later pushed to lower and lower wavelengths (down to $4.2\,nm$) at the $1.2\,GeV$ FLASH facility at DESY, in Hamburg [12], [13]; the 2009 results [14] at the $14\,GeV$ Linac Coherent Light Source (LCLS) at SLAC in Stanford, California, demonstrated SASE lasing at $0.15\,nm$, opening the era of hard x-ray FELs. In spring 2011, SASE amplification at $0.12\,nm$ was also observed at the $8\,GeV$SACLA facility at SPring-8 in Japan [15], and more recently down to $0.08\,nm$.

The linear accelerator (linac) geometry is essential in allowing the low emittance and the high peak current required to trigger the SASE process. During acceleration in a linac, the normalized emittance $\epsilon_n = \gamma\epsilon$ is approximately constant, and this implies that the emittance decreases as the energy γ grows. So, if a sufficiently low emittance is available already at the start, i.e. at the electron gun of the injector system, emittances well below the $0.1\,nm$ range are achievable. Today's state of the art injector systems achieve normalized emittances of order or below $1\,\mu m\,rad$ for bunch charges of order $1\,nC$, and even lower for reduced bunch charges. Since $\gamma \simeq 20,000$ at an energy of $10\,GeV$, emittances below $\epsilon \simeq 0.05\,nm\,rad$ are achievable. Furthermore, the required high peak currents can be achieved by compressing the bunches in one or several suitable magnetic chicanes, down to bunch lengths of order $\simeq 10\mu m$, corresponding to pulse durations of $30\,fs$. With lower bunch charges, few fs pulses were demonstrated at LCLS [16].

It is important to underline that, whereas in a synchrotron source the limitation on attainable wavelengths is only related to the tunability of undulators (i.e. to the possibility of modifying the K parameter by using a mechanical movement of the magnet support structure to change the magnetic field in the device), in a SASE FEL the achievement of saturation requires a sufficient

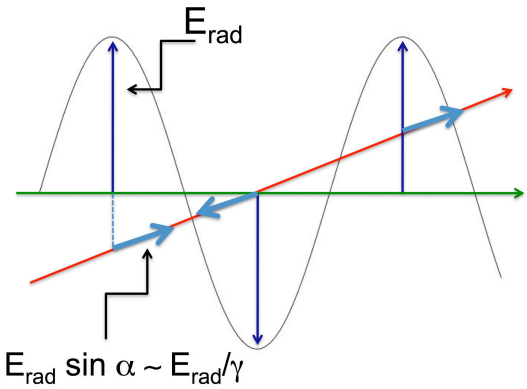


Fig. 2: *Schematics of the microbunching process. The radiated electromagnetic wave (black wavy line) propagates along the undulator axis (green line), and the electron trajectory (red line) is at an angle $\alpha \sim 1/\gamma$ to this axis. Therefore the radiation electric field E_{rad} has a small component (blue arrows) parallel or antiparallel to the electron velocity, $E_{rad}sin\alpha$, which can perform work on the electrons and therefore accelerate or decelerate them.*

length of the undulator. The characteristic length of the exponential intensity growth is the gain length, L_g, which is roughly proportional [17] to ${\epsilon_n}^{5/6}/I^{1/2}$; here the normalized emittance and the peak current of the bunch are the crucial quantities. A good estimate for the saturation length is a factor 9 or 10 times the gain length. Therefore, for SASE lasing at a given wavelength it is not only necessary to tune the undulator at that wavelength; the emittance must be small enough and the peak current high enough to ensure that the saturation length is shorter than the undulator length.

3 Scientific case for hard X-ray Free-Electron Lasers

The most important features of the x-ray pulses of Free-Electron Lasers (FEL's) are the short duration, typically on a few $10\,fs$ time scale, the high peak brilliance, translating into a number of photons per pulse in the $10^{11} - 10^{12}$ range; the very high degree of transverse coherence [18]; and a typical bandwidth of the pulses in the $\Delta\lambda/\lambda \simeq 10^{-3}$ range. This means that the number of photons, that typically reach the sample in one second in an experiment on the best synchrotron beamlines, can be delivered on some $\simeq 30\,fs$ in an FEL experiment.

The very short duration of pulses and the high degree of coherence are beginning to deliver a big scientific payoff in x-ray structural experiments. In traditional crystallography, x-ray diffraction is used to unveil the electron density, for example in a molecule, by analyzing the intensity distribution of the Bragg peaks in the diffraction pattern of a crystal of the molecule in question. This is done because the signal is enhanced by the coherent superposition process at the origin of Bragg reflections (scattering power growing with $\simeq N^2$ rather than N = number of molecules) and also because the large number of photons needed to acquire the signal is distributed between many molecules, limiting the effects of radiation damage. On the other hand, the spatial periodicity of the crystal is not an indispensable route to the acquisition of sufficient information to reconstruct the electronic density of a system; theoretical [19] and experimental [20] evidence show that collection of data with a spatially coherent soure, on a fine grid of scattering vectors (more precisely, "oversampling" on a grid finer than the Nyquist spacing, i.e. the inverse of the size of the diffracting specimen) can allow the solution of the phase problem by iterative algorithms. This applies to both general non-periodic objects and to very small crystals, smaller than the illuminated volume, in which periodicity is broken by the sample surfaces. In principle, delivering some $10^{11} - 10^{12}$ photons to such a sample, reconstruction is possible. However the problem of radiation damage imposes very tight constraints, especially for biological samples. Photoelectron and Auger electron emission induces a high number of defects, charging of the target and raising of its effective temperature, with a large number of atoms displaced from their original positions. Here the short duration of the FEL pulse, however, brings a decisive advantage: data collection takes place on a time scale too fast for the atoms to move, the observed structure is therefore the unperturbed one, even if the sample is deeply distorted, or even completely destroyed in the process.

The first experimental demonstration of this principle was the single-pulse coherent diffraction imaging by Chapman *et al*. [21], in which a diffraction pattern sufficient to reconstruct an image by standard iterative algorithms was acquired using a single FLASH pulse of $25\,fs$ duration. As a result of the high number of photons in the pulse, photoelectric absorption deposits sufficient energy in the sample (a microstruture milled through a silicon nitride membrane) to bring it up to a temperature of 60,000 K and to destroy it completely. Nonetheless, the extremely short duration of the pulse allows collection of the relevant data before the sample is blown apart.

Although the use of 32 nm radiation limits the resolution to a few tens of nm, which could be obtained easily by other, non-destructive, methods, the interest of this experimental breakthrough lies in the proof-of-principle of single-shot imaging of non-periodic objects. One of the chapters of the scientific case for hard x-ray FEL's is the hope to be able to image non-periodic biological objects (from individual cells down to large macromolecules), with resolution approaching the atomic scale, without the need for crystallization - which is a major hurdle in structural biology studies. Very significant steps towards this goal were achieved at the LCLS, where images of single large viruses were acquired and reconstructed [22], and sub-nm resolution structures of biomolecules were obtained from nanocrystalline samples, dispersed in aqueous solution [23].

One can imagine other fields in which the possibility to acquire images on an ultra-fast time scale can be important. For example, in the study of liquids with x-rays (and even more so with neutrons) so far, the acquisition time has always been much larger than that of the disordered translational and rotational molecular motions that permanently rearrange the configuration: only average quantities such as radial distributions are measurable. With an FEL source one can take snapshots of instantaneous configurations and think of questions such as the statistics of configurations in a liquid versus that in the amorphous solid; or a real time observation of nucleation phenomena at the liquid-solid boundary.

The possibility of single-shot structural information, as confirmed by these novel experiments, opens naturally the door to the study of time-dependent phenomena on the atomic scale. If one can get structural information on a system in a very short time, one could dream of following the evolution of chemical reactions in time, e.g. biochemical processes, catalytic mechanisms, and so on. However, the sample destruction by a single shot manifestly interrupts the time evolution one would like to investigate. There is a way out, though, as long as the object of investigation (the molecules of the reagents, for example) is available in many indistinguishable copies, and there is a possibility of fast "triggering" of the process; for example, if we consider a photochemical reaction triggered by an IR laser flash, the "pump and probe" experimental strategy can be put to work: we repeat the experiment on many copies of the system, each time enforcing a different time delay between the start of the process and the interrogation by an FEL pulse. Each acquisition is like one snapshot of a movie, and when they are put together, they deliver the time evolution of the phenomenon. There are limits to the precision of determining the time delay of an IR laser pulse and an FEL pulse: but the experience so far acquired at FLASH and LCLS shows that this is possible with an accuracy of the order of one or a few hundreds of fs (see for example [24]). This is still one order of magnitude longer than the duration of either pulse, but it is an interesting time scale for a variety of photochemical processes.

So far only measurements of structural quantities such as the charge density were discussed. We know, on the other hand, that other important observables and order parameters, such as magnetic moments, and other electronic order parameters such as orbital ordering are also accessible to x-ray investigations, especially in the resonant scattering regime. Recent experiments demonstrated the possibility of pump-probe studies of these order parameters (see [24]). Measurement of a single-shot resonant magnetic scattering pattern at the Co M edge of a Co-Pt multilayer system was reported in [25]).

In addition, the remarkable transverse coherence can be used to probe fluctuation dynamics by X-ray Photon Correlation Spectroscopy, with much increased possibilities with respect to those achievable in a synchrotron source.

Other uses of powerful FEL pulses are envisaged in plasma physics and more generally in the

Table 1: *Basic parameters of the three hard x-rays FEL projects (see text); brilliances are expressed in* $photons/s/mrad^2/mm^2/0.1\%BW$.

Project	LCLS	SACLA	European XFEL (SASE1)
Max. Electron Energy (GeV)	14.3	8.0	17.5
Min. Photon Wavelength (nm)	0.15	0.08	0.05
Photons/pulse	$\sim 10^{12}$	2×10^{11}	$\sim 10^{12}$
Peak Brilliance	1.5×10^{33}	1×10^{33}	5×10^{33}
Average Brilliance	4.5×10^{22}	1.5×10^{23}	1.6×10^{25}
Pulses/second	120	60 X N	27 000
Date of first beam	2009	2011	2015

study of high energy-density states of matter. The creation of plasma-like or other highly excited states of matter by FEL pulses has been mentioned before, essentially as a drawback for imaging experiments. However, to achieve in the laboratory conditions usually found only in astrophysical environments, and to investigate portions of the temperature/density phase diagram of materials usually not encountered at the surface of the earth, is very appealing. A vigorous activity in this field already exists, using powerful (up to petawatt) IR or visible lasers to generate the corresponding conditions. X-ray pulses have the definite advantage of a much larger penetration depth, resulting in a more uniform excitation profile in the sample part accessible to the probe (e.g. a subsequent x-ray pulse).

4 The European XFEL and the international competition for hard x-ray FEL's

The remarkable results obtained at the FLASH facility have allowed considerable progress in the understanding of the SASE process itself, and demonstrated the revolutionary potential of FEL experiments for a variety of disciplines. This has provided further stimulation to projects for the realization of hard x-ray FEL's. There are at present three major projects worldwide, one in the USA (the Linac Coherent Light Source, LCLS, in Stanford, California [26]) which obtained the first beam at 0.15 nm in April 2009; one in Japan (the SACLA, SPring-8 Angstrom Compact Laser, at SPring-8 [27]), currently completing its commissioning and expecting the first users soon; and one in Europe (the European XFEL in Hamburg, currently under construction [28]); all of them target wavelengths of the order or smaller than $0.1\,nm$ suitable for experiments determining structural properties with atomic resolution. The main features of the three projects are summarized in Table I.
More recently projects in Switzerland [29] and in South Korea [30] were started.
The LCLS project in Stanford has been welcoming users for experiments since late 2009. It uses the pre-existing SLAC high energy linear accelerator, or, more precisely, one third of its length, to accelerate electrons and feed them into an undulator with fixed gap, $112\,m$ long, to produce coherent x-rays with photon energies between 0.8 and 8 keV, i.e. with wavelengths between 1.5 and $0.15\,nm$. Tuning of the photon energy occurs by tuning of the electron energy

Fig. 3: *Layout of the European XFEL, showing the trace of the underground tunnels running under the northwest districts of Hamburg and the neighbouring town of Schenefeld*

between 4.5 GeV and the maximum energy 14.3 GeV. The linear accelerator has been modified to accept the bunches produced by a RF photocathode gun and preserve the low-emittance beams required for the SASE process; and also to include two stages of bunch compression. The repetition rate is 120 pulses per second. Remarkably, the low emittance values obtained by lowering the charge of individual bunches allowed achievement of robust SASE lasing [14] at $0.15\,nm$, with attainment of saturation intensity requiring only about one half of the available undulator length. These very exciting results provide strong confirmation to the validity of the SASE principle and to the reliability of computational simulations of the process.

The Japanese SACLA project is characterized by the attempt to reduce the size and the cost of large FEL installations by daring innovations, such as the use of a thermionic cathode gun to produce very low emittance bunches, the use of C-band accelerator technology to generate very high acceleration gradients, a very high compression ratio for the bunch length (4,000) and the use of a tunable in-vacuum undulator, with a $4\,mm$ gap in the standard operation for generation of $0.1\,nm$ radiation with an electron energy of $8GeV$, considerably lower than in the competing projects. The design value for the number of pulses per second varies, as there are 60 RF pulses per second, and each of them can be filled with more than one bunch. There are five undulators foreseen in the ultimate configuration of the facility. So far the commissioning is making good progress with lasing down to $0.8\,nm$ reported.

The European X-ray Free-Electron Laser Facility (European XFEL), which started construction in January 2009 in Hamburg, is deriving its basic technical choices from the successful FLASH experience. The photoinjector and the gun are directly derived from the corresponding FLASH components and so is the basic superconducting accelerator technology, first developed in the context of the international TESLA collaboration, coordinated by DESY. The 1.7 km long accelerator, located in an underground tunnel (see Fig. 3), can provide electrons up to $17.5\,GeV$

energy, and feed them into two beamlines according to the scheme displayed in Fig. 4 [28]. The first beamline contains a hard x-ray undulator (SASE1), for $\leq 0.1\,nm$ coherent photons and a soft x-ray one (SASE3), which can sometimes make use of the "spent" beam resulting from saturation of SASE1 to generate soft X-rays in the 0.4 to 1.6 nm range (at 17.5 GeV electron energy: softer x-ray radiation is of course obtained if the electron energy is reduced). The second beamline contains a second hard x-ray undulator (SASE2), identical to SASE1, and two tunnels downstream in which two further undulators can be located. In the baseline design for the initial phase of the facility, each of the three SASE undulators will feed into two instruments. The six instruments foreseen to become available in 2015 are:
Hard x-ray instruments: 1. Materials Imaging and Dynamics (MID); its purpose is to use the coherence of the beams to explore dynamic fluctuations in matter, accessing unprecedented length and time scales.
2. X-ray Femtosecond Experiments (FXE) the purpose of this instrument is to explore ultrafast phenomena in the physics and chemistry of solids, liquids and soft-matter systems.
3. Single Particle, Cluster and Biomolecular Imaging (SPB); this instrument is devoted to the pursuit of structural studies in non-periodic systems, especially in structural biology.
4. High Energy-Density Science (HED); here the idea is to use the FEL pulses to bring a target to extremely high values of temperature and then to interrogate it in order to access regions of the phase diagram not easily accessible in the laboratory (e.g. warm dense matter, in which ordinary densities of solid materials are present at temperatures such that $k_B T \simeq 1 - 10eV$).

The soft x-rays instruments are:
1. Small Quantum Systems (SQS); this instrument should continue and extend to higher photon energy the innovative experiments pioneered at FLASH on ions, atoms and molecules.
2. Soft X-ray Coherent Scattering and Spectroscopy (SCS); this is a multi-purpose instrument, on which different end-stations should be mounted for a variety of techniques in the soft x-ray analysis of materials, such as RIXS, absorption, photoemission, etc.

The use of the superconducting technology is a formidable advantage of the European facility: it allows a very wide flexibility in the operating conditions; in particular, it allows to fill each RF pulse with a very large number of electron bunches. In the European XFEL it is foreseen to have a train of up to 2,700 bunches in each of the 10 RF pulses (of 600 μs duration) per second. It will be possible to switch the electrons from one beamline to the other during each bunch train. The possibility to use such a large number of bunches, with a spacing of 220 ns, implies considerable development work in the field of detectors, as well as in the lasers for pump-probe experiments, which should be able to follow the time structure of the XFEL pulses. This is already in progress. The average brilliance, corresponding to this large number of X-ray pulses per unit time, could prove very important in experiments such as coherent diffraction of non-periodic objects, where hits of the FEL pulses with a molecule are expected to be very rare, but a large number of them needs to be accumulated in order to achieve a satisfactory signal-to-noise ratio.
The civil construction started in January 2009, with advance funding from the German government. At present, a Convention concerning the Construction and Operation of a European X-ray Free-Electron Laser Facility was signed by representatives of twelve countries (Denmark, France, Germany, Greece, Hungary, Italy, Poland, Russia, Slovakia, Spain, Sweden, Switzerland), The largest contributors to the construction costs are Germany and Russia. The Convention foresees the creation of a limited liability company under German law, which exists, under

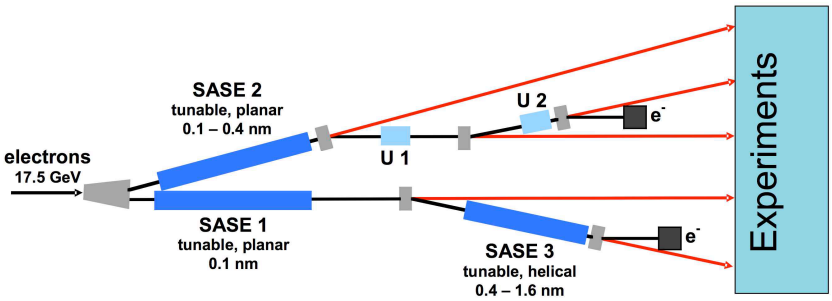


Fig. 4: *Schematic layout of the arrangement of undulators at the European XFEL. The initial configuration of the facility does not include the spontaneous emission undulators U1 and U2, nor the helical version of SASE3, which is going to be initially built in the planar version.*

the name European X-ray Free-Electron Laser Facility GmbH, since September 2009, and is in charge of the project.

5 Future perspectives: self-seeding of hard x-ray FEL's

We have so far mentioned the transverse coherence properties of SASE FEL radiation, but did not address the issue of its longitudinal (or temporal) coherence. In point of fact, longitudinal coherence is very poor, and a SASE x-ray pulse is composed of a very large number of incoherent spikes (see Fig. 5). The reason can be understood as follows. In order to achieve full longitudinal or temporal coherence the microbunching described in Section 2 should uniformly encompass the whole bunch length: a density modulation should extend from the tail to the head of the bunch without interruption or phase variation. This cannot happen in most cases because the microbunching is imprinted by the emitted radiation as it catches up with the electrons that were ahead of the emission point. But the resonance condition for the undulator (Eq. (1) with $n = 1$) ensures that the radiation gains a wavelength λ over the electrons in a distance λ_u (a magnetic period). This implies that in an undulator with N magnetic periods, the maximum length of a coherent microbunching pattern is $N\lambda$. If this is shorter than the bunch length, one can expect different parts of the bunch to display unrelated (phase incoherent) microbunching, each producing a SASE spike with no phase relationship to the others. A coarse estimate of the number of spikes is, according to the previous argument $N_s = L_b/N\lambda$, where L_b denotes the bunch length. For typical parameters for the European XFEL, where $\simeq 4000$ undulator periods of $40\,mm$ constitute the SASE1 or SASE2 undulators, and a bunch length of $30\,\mu m$, corresponding to $100\,fs$ duration, at $0.1\,nm$ one can estimate $N_s \simeq 70$. On the same basis the duration of a spike is of order of $1\,fs$ and this the basic estimate of the coherence time of the source, which is very small.

In order to improve the longitudinal coherence, that is to obtain a smooth single-mode lineshape instead of the ragged pattern of Fig. 5 , various *seeding* schemes are possible. The idea behind seeding is to start the amplification process not by random shot noise fluctuations in the bunch electron density distribution, but by an external better controlled radiation pulse, with a pulse energy exceeding that of random fluctuations. In longer wavelength FEL's the external pulse is provided by an IR or visible laser, that is used to generate suitable harmonics, either in a non-linear medium, or in an additional undulator. This is the principle of seeding schemes, as tested in Japan [31] and adopted in soft x-ray facilities such as FERMI@Elettra in Trieste [32].

In hard x-ray facilities, generation of laser harmonics is not going to work, as the order of the harmonic would be too high. Therefore various versions of *self-seeding* are being thought of, in which the seeding radiation is produced in a short undulator with the same parameters as the long SASE one that follows. The radiation from the first undulator is monochromatized and superposed again to the electron bunch in the SASE undulator. The most promising *self-seeding* scheme was proposed in [33] and should be tested at LCLS. A self-seeded x-ray FEL could have a $\Delta\lambda/\lambda$ between 10^{-4} and 10^{-5} , with an enhancement of the peak power in the $\simeq 100GW$ range and would be a remarkable progress in the field.

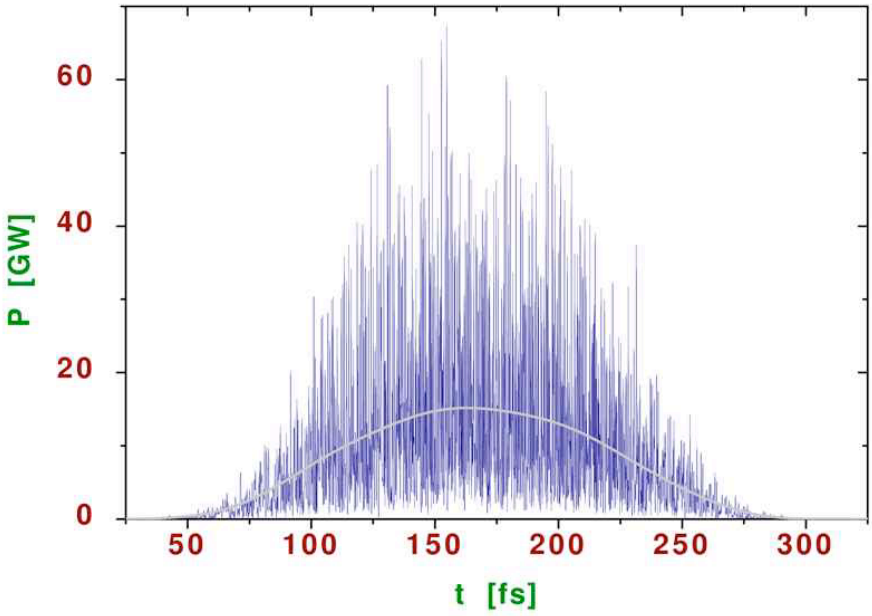


Fig. 5: *Simulation of the temporal profile of a 12.4 keV European XFEL pulse from SASE1 (Courtesy of M.V. Yurkov)*

6 Conclusions

As we tried to illustrate briefly, this is a very exciting time in the development of accelerator-based light sources, as many new and revolutionary facilities are starting operation. Their promise is great, and the scientific user community should accordingly prepare for them. The scientific pay-off of the new facilities, in fact, will to a large extent be determined by the progress of instrumentation for experiments. The new extremely bright and ultrafast sources require, for their full exploitation, corresponding progress in optics, diagnostics, and above all detectors and data acquisition strategies. They might also imply a change in the size and composition of the experimental teams. The early attention to instrumental issues and a close relationship with potential users from the very beginning may turn out to be a factor of importance for the success of the new generation of light sources.

References

[1] D.H. Bilderback, P. Elleaume, E. Weckert, J. Phys. B: At. Mol. Opt. Phys. **38**, S773 (2005)

[2] K.Balewski *et al*., *PETRA III: a low emittance synchrotron radiation source, Technical Design Report*, DESY 2004-035, (2004)

[3] For a recent excellent review, see M. Bei *et al*., Nucl. Instr. Methods A**622**, 518 (2010)

[4] J.Ablett *et al*.,*NSLS-II,Conceptual Design Report*, (http://www.bnl.gov/nsls2/project/CDR/), 2006

[5] S.C. Leemann *et al*., Phys. Rev. ST AB **12**, 120701 (2009)

[6] R. Hettel *et al*., *Concepts for the PEP-X light source*, Proceedings of the 2009 Particle Accelerator Conference, http://accelconf.web.cern.ch/AccelConf/PAC2009/papers/we5rfp015.pdf

[7] R.W. Schoenlein *et al*., Science **287**, 2237 (2000)

[8] A.M. Kondratenko, E.L. Saldin, Sov. Phys. Doklady **24**, 986 (1979)

[9] R. Bonifacio, C. Pellegrini, L.M. Narducci, Opt. Commun. **50**, 373 (1984)

[10] J.B. Murphy, C. Pellegrini, Nuclear Instrum. Methods A**238**, 159 (1985)

[11] S.V. Milton et al., Science **292**, 2037 (2001)

[12] V. Ayvazyan et al., Phys. Rev. Lett. **88**, 104802 (2002)

[13] M. Vogt *et al*., *Proceedings of IPAC2011, San Sebastin, Spain*, http://accelconf.web.cern.ch/AccelConf/IPAC2011/papers/thpc081.pdf

[14] P. Emma, *Proceedings of PAC09, Vancouver, BC (Canada)*, http://trshare.triumf.ca/pac09proc/Proceedings/papers/th3pbi01.pdf

[15] H. Tanaka, *Proceedings of IPAC2011, San Sebastin, Spain*, http://accelconf.web.cern.ch/AccelConf/IPAC2011/papers/moyca01.pdf

[16] Z. Huang *et al*., *Proceedings of 2011 Particle Accelerator Conference, New York, NY, USA*, http://accelconf.web.cern.ch/AccelConf/PAC2011/papers/thp183.pdf

[17] E. Saldin, E.A. Schneidmiller and M.V. Yurkov, Optics Commun. **235**, 415 (2004)

[18] I. Vartanyants *et al*., Phys. Rev. Lett. **107**, 144801 (2011)

[19] D. Sayre, in *Imaging Processes and Coherence in Physics*, M. Schlenker *et al*. eds., 229235 (Springer, Berlin, 1980).

[20] J. Miao, P. Charalambous, J. Kirz and D. Sayre, Nature **400**, 342 (1999)

[21] H.N.Chapman *et al*., Nature Phys. **2**, 839 (2006)

[22] M.M. Seibert *et al*., Nature **470**, 78 (2011)

[23] H.N. Chapman *et al*., Nature **470**, 73 (2011)

[24] N. Pontius *et al*., Applied Phys. Lett. **98**, 182504 (2011)

[25] C. Gutt *et al*., Phys. Rev. B**81**, 100401(R) (2010)

[26] For updated information, see http://lcls.slac.stanford.edu/

[27] For updated information, see http://xfel.riken.jp/eng/

[28] M. Altarelli *et al*., ed.s, *The European X-ray Free-Electron Laser: Technical Design Report*, http://www.xfel.eu/en/documents/

[29] For updated information, see http://www.psi.ch/swissfel/swissfel

[30] H.S. Kang and S.H.Nam, in *Proceedings of FEL2010, Malmö, Sweden*, 155 (2010), http://fel2010.maxlab.lu.se/FEL2010_proceedings/papers/mopc19.pdf

[31] G. Lambert *et al*., Nature Phys. **4**, 296 (2008)

[32] see for example E. Allaria *et al*., New J. Phys. **12** 075002 (2010)

[33] G. Geloni, V.Kocharyan, E. Saldin, Journal of Modern Optics **58**, 1391 (2010)

C 9 The road ahead for the European Spallation Source

C. J. Carlile
European Spallation Source Scandinavia, Director
Lund University, Sweden

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

With the political agreement in early summer 2009, to choose a site for the ESS, 17 European countries have joined together in a partnership to embark on a 3-year Pre-Construction Phase, including a full Design Update and Preparation to Build activities, prior to formally starting construction on the site in Lund in southern Scandinavia in 2013.

The ESS will be a 5MW long-pulse spallation neutron source and as such it will be a unique and uniquely powerful facility offering new opportunities for materials research using slow neutrons. Compared to the user's experience today the ESS aims to offer measuring capabilities two orders of magnitude better than similar facilities today.

Our accelerator will be substantially superconducting using niobium for the cavities. Interestingly, forty years ago, in my PhD thesis work and subsequently our research group at Birmingham were investigating the diffusion of hydrogen in palladium and niobium. Such metals suck up hydrogen from the environment and it is interesting that today the absorption of hydrogen by niobium is very relevant in accelerator component technology. A large effort in understanding hydrogen in metals was made in Jülich in the 1970s and 1980s but as with much work at that time cross-fertilisation was not frequent and the solid state and material sciences community had little inter action with the accelerator community. That is much less the case nowadays.

Planning for a new 1.5 B€ scientific facility which will be operational a decade from now requires a new approach. For example it is likely by 2020 that the way in which researchers interact with central science facilities will have changed dramatically thanks to the very rapid advance of computing and IT methods. Equally well, the energy consumption of big facilities puts a significant burden on the annual operating budget of such facilities and novel methods must be found to manage the energy inventory of the ESS. One consequence of the high demand for energy is the environmental impact, which has political overtones.

ESS will be built on a truly green field site. Green field in the psychological as well as the physical sense. This gives an opportunity to revisit the standard methods of dealing with the user community, providing facilities, which will give added value to the visiting researcher and those who remain at their home laboratories. Thanks to the construction of the high brightness synchrotron source MAX IV being built in the same location as ESS mutual advantages by bringing the user communities of the two sources together will be a gain for all. A seventeen hectare piece of land between the two facilities will be used to build INXS, the institute for neutron and x-ray science which will incorporate facilities for users such as meeting places and conference rooms but, more importantly, an assembly of separately identifiable laboratories similar in nature to the Partnership for Structural Biology which was so innovative on the ILL-ESRF site in Grenoble. Similar centers for soft condensed matter, material under extreme conditions and material for climate, energy and health will be explored. I will attempt to address these questions and give a personal view of future developments. Further information can be obtained on our website www.esss.se , on Wikipedia, and on the NMI3 site http://nmi3.eu/

D 1 Small Angle Scattering and Large Scale Structures[1]

H. Frielinghaus
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 43rd IFF Spring School “Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources” (Forschungszentrum Jülich, 2012).

1 Introduction

Scattering methods have to be seen in context to direct imaging methods (such as microscopy). The wavelength of the probe gives a good guess for the accessible structural sizes to be resolved – in either case it is connected to a lower boundary. Especially for the probes neutrons and x-rays, small angle scattering opens a window on the nanometer length scales. These length scales are highly interesting as mesoscales where the atomistic properties are overcome and many body effects come into play. A lot of macroscopic properties can only be explained on the basis of mesoscopic length scale effects. Theoretical understanding, modeling and small angle scattering are tightly linked.

Historically, lens systems for neutrons and x-rays were not accessible and, thus, the only way to the nanometer length scales were scattering experiments [1, 2]. While direct images are intuitively understandable, scattering experiments have to be understood on the basis of the Fourier transformation. While for well ordered crystals the understanding remains a simple task (note Bragg's law) for many mesoscopic effects the scattering patterns demand for more complicated modeling. Nonetheless, scattering models have been and are developed for the focus of the actual research, and apart from that, plausible arguments are found for many practical applications such that the experienced small angle scatterer sees a lot of details at first hand before tedious modeling is applied.

Nowadays, electron microscopy has become a strong tool to probe the nano-world. It has to be mentioned, that often surfaces (of fractures) are characterized. Single images are only excerpts of the whole story – sometimes one even has to worry whether the conditions of preparation represent the desired condition. Apart from that, average magnitudes – such as particle sizes – have to be counted from many measurements. The scattering experiments are in most of the cases in situ experiments. The sample can be kept at a certain temperature, pressure, magnetic and electric field – even shear and flow fields can be applied. In most of the cases, rather large volumes are irradiated, and so an average impression is obtained from macroscopic volumes. So, for instance, the obtained particle sizes base on good statistics.

A good example for small angle scattering are polymer systems. A polymer is a linear molecule of identical monomeric units – so polystyrene is made of the styrene monomer. The monomer properties are to be understood (more or less) as atomistic effects, and mainly the connectivity of the monomers is responsible for the mesoscopic understanding which finally ends in a coil-shaped linear molecule. The small angle scattering experiment especially focuses on the connectivity and is able to resolve the overall dimension of the single molecule. If chemically different polymers form a compound material, often nanometer domains form – so the polymers tend to phase separation. These domains can be tailored by the thermal history, but also by chemical additives (such as diblock copolymers) which support the miscibility by keeping the domains small at a desired size. Still, these length scales are of several tens to hundreds of nanometers and are ideally observed by small angle scattering. The mixing of polymers leads to mixed macroscopic properties – so often mechanic properties are tailored from originally individual favorable properties such as softness, rubber elasticity, and mechanical strength. A microscopic understanding of these effects is invaluable.

Complex fluids consist of at least two components. In the case of microemulsions oil and water are mixed macroscopically homogenously – but microscopically still domains of oil and water form on the nano-scale. The added surfactant makes the microemulsion thermodynamically stable and controls the domain size. Additives like amphiphilic polymers tailor the overall properties of the amphiphile. Often, the surfactant film becomes stiffer, and larger domains can

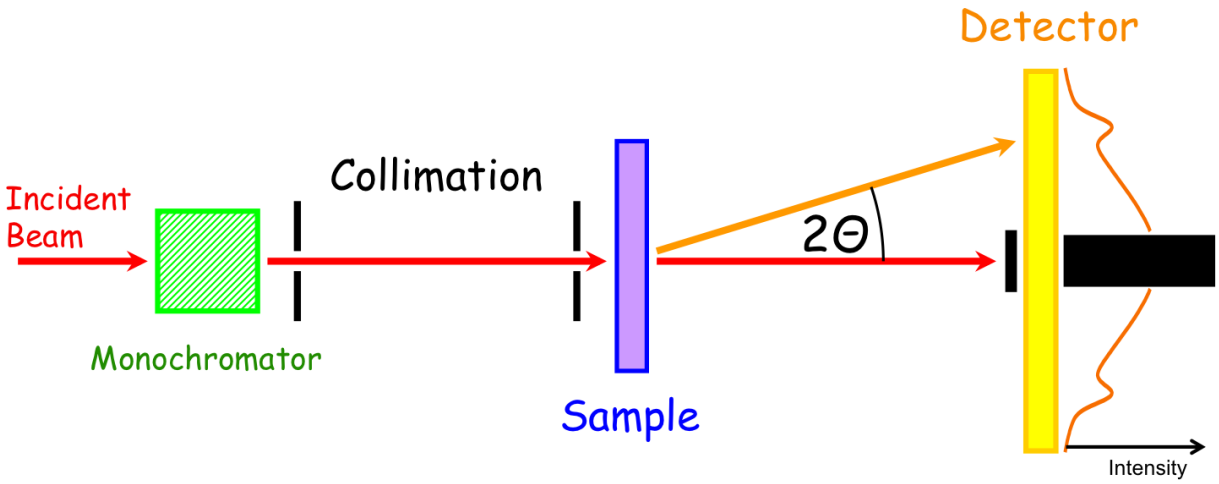


Fig. 1: *Scheme of a small angle neutron scattering instrument. The neutrons pass from the left to the right. The incident beam is monochromated and collimated before it hits the sample. Non-scattered neutrons are absorbed by the beam stop in the center of the detector. The scattered neutron intensity is detected as a function of the scattering angle* 2θ.

form. All these effects were monitored by small angle neutron scattering experiments.

2 Small Angle Neutron Scattering

The first small angle neutron scattering instruments have been developed for continuous sources realized by research reactors. The neutrons are moderated by the cold source to ca. 20-30K. Typical wavelengths are found between 4 and 6Å, and reasonable intensities are obtained up to 20Å. A typical scheme of a small angle neutron scattering (SANS) instrument is depicted in Figure 1. The cold neutrons can easily be guided by mirror coated cavities to the instrument. The first important element is the neutron velocity selector (Fig. 2). It consists of tilted lamellae on a rotating cylinder which selects the desired neutron velocity mechanically. The wavelength uncertainty results from the gap width and length between the lamellae and typically takes values of ± 5 to $\pm 10\%$. Then the collimation is formed by an entrance aperture and the sample aperture with a certain distance L_C on which the neutrons propagate freely. For varying the collimation distance, the neutrons are guided with movable 1m neutron guides to the desired distance relatively to the fixed sample position. Very often, the maximum length is 20 meters, but other examples exist with 6 over 10 to 40m. The sample is placed directly behind the sample diaphragm (see sample position on Fig. 3). Most of the neutrons (between 50 and 90%) pass the sample unscattered. In front of the detector a beam stop absorbs these high intensities because the highly sensitive detector does not serve this high dynamic range. The scattered neutrons are detected on a position sensitive area detector. In this way, the scattering intensity is monitored as a function of the scattering angle 2θ. Typical sizes of the beam stop are $4\times 4\text{cm}^2$, and for the sensitive area between 60×60 to $100\times 100\text{cm}^2$. For varying the covered angle, the detector is placed at different distances L_D (symmetrically to the collimation distance L_C) inside a large vacuum tube.

Fig. 2: *The neutron velocity selector of the small angle scattering instrument KWS-3 at the research reactor Garching FRM-2. This selector was especially manufactured for larger wavelengths (above 7Å).*

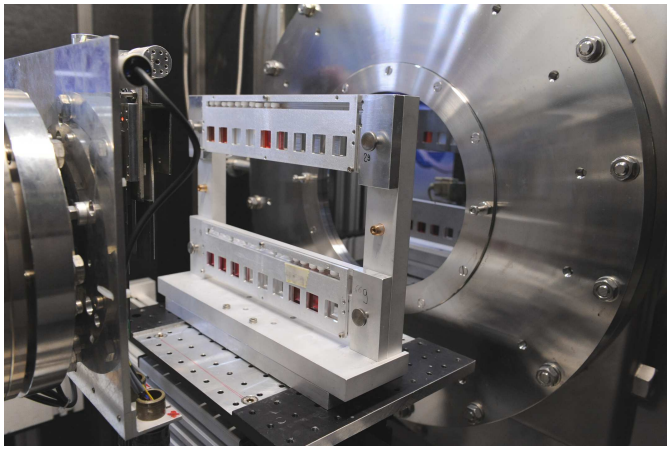


Fig. 3: *View on the sample position of the small angle scattering instrument KWS-1 at the research reactor Garching FRM-2. The neutrons come from the left through the collimation and sample aperture (latter indicated). A sample changer allows for running 27 samples (partially colored solutions) in one batch file. The silicon window to the detector tube is seen behind.*

2.1 The scattering vector Q

In this section, the scattering vector $\mathbf{Q}$ is described with its experimental uncertainty. The scattering process is schematically shown in Fig. 4, in real space and momentum space. In real space the beam hits the sample with a distribution of velocities (magnitude and direction). The neutron speed is connected to the wavelength, whose distribution is depending on the velocity selector. The directional distribution is defined by the collimation. After the scattering process, the direction of the neutron is changed, but the principal inaccuracy remains the same. The scattering angle 2θ is the azimuth angle. The remaining polar angle is not discussed further here. For samples with no preferred direction the scattering is isotropic and, thus, does not depend on the polar angle. In reciprocal space the neutrons are defined by the wave vector $\mathbf{k}$. The main direction of the incident beam is defined as the z-direction, and the modulus is determined by the wavelength, so $|\mathbf{k}_i| = 2\pi/\lambda$. Again, $\mathbf{k}$ is distributed due to the selector and the collimation inaccuracies. The wave vector of the (quasi) elastic scattering process has the same modulus, but differs in direction, namely by the angle 2θ. The difference between both wave vectors is given by the following value:

$$Q = \frac{4\pi}{\lambda} \sin\theta \tag{1}$$

For isotropic scattering samples, the measured intensity depends only on the absolute value of the scattering vector $Q = |\mathbf{Q}|$. For small angles, the common approximation of small angle (neutron) scattering is valid:

$$Q = \frac{2\pi}{\lambda} \cdot 2\theta \tag{2}$$

The typical Q-range of a small angle scattering instrument thus follows from the geometry. The detector distances L_D vary in the range from 1m to 20m. The area detector is active between

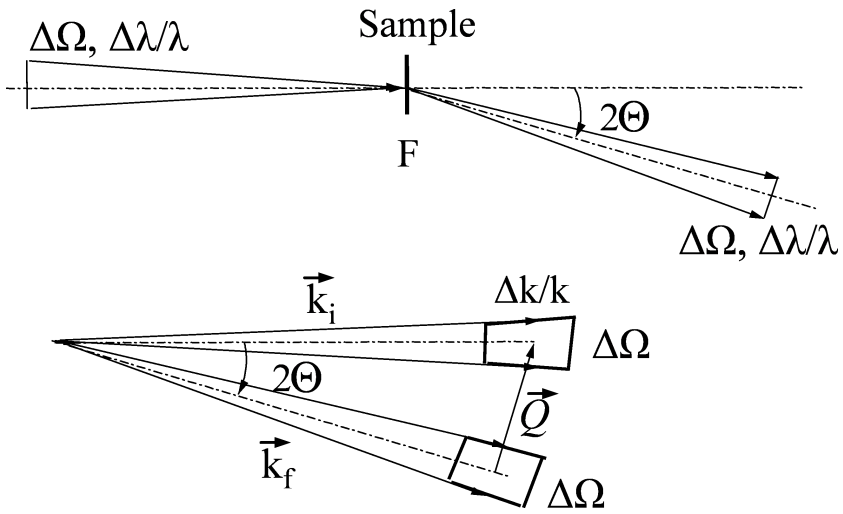


Fig. 4: Top: *the neutron speed and its distribution in real space, before and after the scattering process.* Bottom: *The same image expressed by wave vectors (reciprocal space). The scattering vector is the difference between the outgoing and incoming wave vector.*

$\varnothing_D = 2$cm and 35cm from the center. The angle 2θ is approximated by the ratio $\varnothing_D / L_D$ and the wavelength λ is 7Å. For the two SANS instruments KWS-1 and KWS-2 at the research reactor Garching FRM-2, a typical Q-range from 1×10^{-3} to 0.3Å^{-1} is obtained.
The Q-vector describes which length scales ℓ are observed, following the rule $\ell = 2\pi/Q$. If a Bragg peak is observed, the lattice parameters can be taken directly from the position of the peak. If the scattering shows a sudden change at a certain Q-value, we obtain the length scale of the structural differences. There are characteristic scattering behaviors that can be described by so called scattering laws that are simple power laws Q^α with different exponents α.

2.2 The Fourier transformation in the Born approximation

This section deals with the physical explanation for the appearance of the Fourier transformation in the Born approximation. In simple words, in a scattering experiment one observes the intensity as the quadrature of the Fourier amplitudes of the sample structure. This is considerably different from microscopy where a direct image of the sample structure is obtained. So the central question is: Where does the Fourier transformation come from?
The classical SANS instruments are also called pin-hole instruments. Historically, pin-hole cameras were discovered as the first cameras. They allowed to picture real sceneries on blank screens – maybe at different size, but the image resembled the original picture. The components of this imaging process are depicted in Fig. 5. Let's assume the following takes place with only one wavelength of light. The original image is then a monochromatic picture of the three numbers 1, 2 and 3. The corresponding rays meet in the pin-hole, and divide afterwards. On the screen, the picture is obtained as a real-space image, just appearing upside down. From experience we know, that the screen may be placed at different distances resulting in different sizes of the image. The restriction of the three beams through the pin-hole holds for the right space behind the pin-hole. In front of the pin-hole the light propagates also in other directions – it is just absorbed by the wall with the pin-hole.
So far, we would think that nothing special has happened during this process of reproduction. But what did happen to the light in the tiny pin-hole? We should assume that the size of the pin-

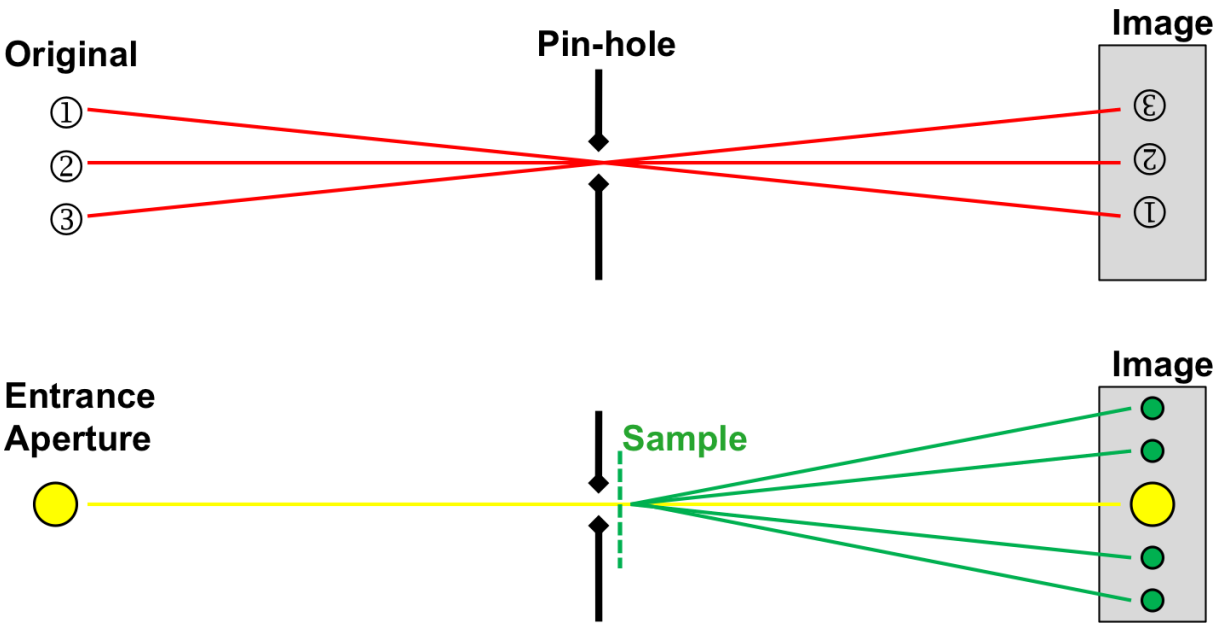


Fig. 5: *The principle of a pin-hole camera transferred to the pin-hole SANS instrument.* Top: *The pin-hole camera depicts the original image (here consisting of three numbers). For simplicity, the three points are represented by three rays which meet in the pin-hole, and divide afterwards. On the screen, a real space image is obtained (upside down).* Bottom: *The pin-hole SANS instrument consists of an entrance aperture which is depicted on the detector through the pin-hole (same principle as above). The sample leads to scattering. The scattered beams are shown in green.*

hole is considerably larger than the wavelength. Here, the different rays of the original image interfere and inside the pin-hole a wave field is formed. The momentum along the optical z-axis indicates the propagation direction and is not very interesting (because is nearly constant for all considered rays). The momenta in the x-y-plane are much smaller and indicate a direction. They originate from the original picture and remain constant during the whole process. Before and after the pin-hole the rays are separated and the directions are connected to a real-space image. In the pin-hole itself the waves interfere and the wave field looks more complicated. The information about the original scenery is conserved through all the stages. That means that also the wave field inside the pin-hole is directly connected to the original picture.
From quantum mechanics (and optics) we know that the vector of momentum is connected to a wave vector. This relation describes how the waves inside the pin-hole are connected to a spectrum of momenta. In classical quantum mechanics (for neutrons) a simple Fourier transformation describes how a wave field in real space (pin-hole state) is connected with a wave field in momentum space (separated beams). In principle, the interpretation is reversible. For electromagnetic fields (for x-rays) the concept has to be transferred to particles without mass. Overall, this experiment describes how the different states appear and how they are related. The free propagation of a wave field inside a small volume (pin-hole) leads to a separation of different rays accordingly to their momentum.
Now we exchange the original image by a single source (see yellow spot in lower part of Fig. 5). This source is still depicted on the image plate (or detector). If we insert a sample at the position of the pin-hole the wave field starts to interact with the sample. In a simplified way we can say that a small fraction of the wave field takes the real space structure of the sample

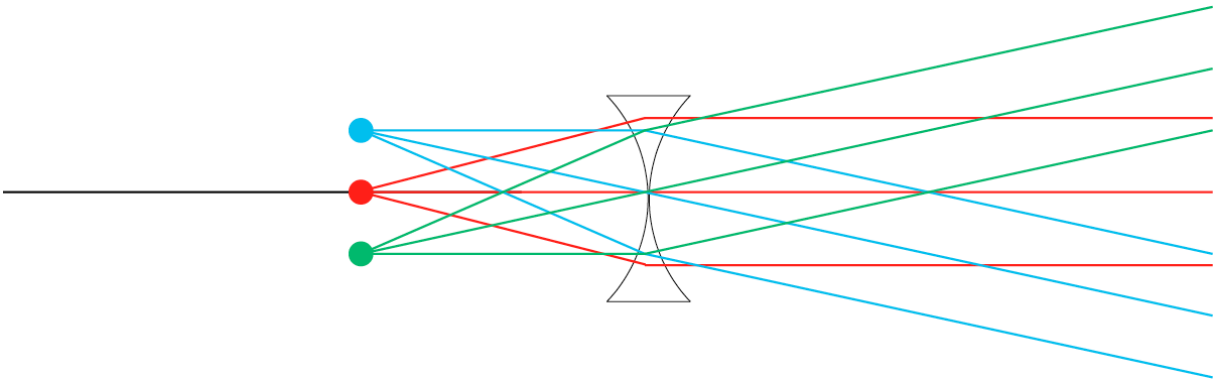


Fig. 6: *How a Fourier transformation is obtained with refractive lenses. The real space structure in the focus of the lens is transferred to differently directed beams. The focusing lens is concave since for neutrons the refractive index is smaller than 1.*

while the major fraction passes the sample without interaction. This small fraction of the wave field resulting from the interaction propagates freely towards the image plate and generates a scattering pattern. As we have learned, the momenta present in the small fraction of the wave field give rise to the separation of single rays. So the real space image of the sample leads to a Fourier transformed image on the detector. This is the explanation how the Fourier transformation appears in a scattering experiment – so this is a simplified motivation for the Born approximation. A similar result was found by Fraunhofer for the diffraction of light at small apertures. Here, the aperture forms the wave field (at the pin-hole) and the far field is connected to the Fourier transformation of the aperture shape.
Later we will see that the size of wave field packages at the pin-hole is given by the coherence volume. The scattering appears independently from such small sub-volumes and is a simple superposition.

2.3 Remarks on focusing SAS instruments

Small angle scattering instruments with focusing optics allow for larger samples. If the entrance aperture remains at the old size the resolution of the experiment is kept, but the intensity is increased tremendously. If the entrance aperture is closed by considerable factors the resolution is increased strongly, and much lower scattering vectors (larger structures) are resolved. Usually, the entrance aperture must have a minimum size for a minimal intensity. For a symmetric set-up (collimation and detector distance equal, i.e. $L_C = L_D$) the focusing optic is in the middle at the sample position. The focus f is half the collimation distance, i.e. $f = \frac{1}{2}L_C = \frac{1}{2}L_D$. Now the places where exact Fourier transforms are obtained (from the entrance aperture and from the sample structure) do not agree anymore. The sample is still considered as a small volume and from there the waves propagate freely to the detector, and the already known relation between sample structure and scattering image holds.
For focusing elements, the places of Fourier transformations differ (see Fig. 6). The original structure is placed in the focus, and the resulting distinctive rays are obtained at the other side of the lens in the far field. So for focusing SAS instruments, the places of appearing Fourier transformations for the entrance aperture and the sample structure differ.
The historical development of cameras can be seen in parallel. The first cameras were pin-hole cameras, but when lenses could be manufactured lens cameras replaced the old ones. The direct

advantage was the better light yield being proportional to the lens size. Another effect appeared: The new camera had a depth of focus – so only certain objects were depicted sharply, which was welcomed in the art of photography. The focusing SAS instrument depicts only the entrance aperture, and the focusing is not a difficult task. The higher intensity or the better resolution are the welcomed properties of the focusing SAS instrument.

2.4 The resolution of a SANS instrument

The simple derivatives of equation 2 support a very simple view on the resolution of a small angle neutron scattering experiment. We obtain:

$$\left(\frac{\Delta Q}{Q}\right)^2 = \left(\frac{\Delta\lambda}{\lambda}\right)^2 + \left(\frac{2\Delta\theta}{2\theta}\right)^2 \tag{3}$$

The uncertainty about the Q-vector is a sum about the uncertainty of the wavelength and the angular distribution. Both uncertainties result from the beam preparation, namely from the monochromatization and the collimation. The neutron velocity selector selects a wavelength band of either $\pm$5% or $\pm$10%. The collimation consists of an entrance aperture with a diameter d_C and a sample aperture of a diameter d_S. The distance between them is L_C.
One property of eq. 3 is the changing importance of the two contributions at small and large Q. At small Q the wavelength spread is nearly negligible and the small terms Q and θ dominate the resolution. This also means that the width of the primary beam is exactly the width of the resolution function. More exactly, the primary beam profile describes the resolution function at small Q. Usually, the experimentalist is able to change the resolution at small Q. At large Q the resolution function is dominated by the wavelength uncertainty. So the experimentalist wants to reduce it – if possible – for certain applications. This contribution is also an important issue for time-of-flight SANS instruments at spallation sources. The wavelength uncertainty is determined by the pulse length of the source and cannot be reduced without intensity loss.
A more practical view on the resolution function includes the geometrical contributions explicitly [3]. One obtains:

$$\left(\frac{\sigma_Q}{Q}\right)^2 = \frac{1}{8\ln 2}\left(\left(\frac{\Delta\lambda}{\lambda}\right)^2 + \left(\frac{1}{2\theta}\right)^2 \cdot \left[\left(\frac{d_C}{L_C}\right)^2 + d_S^2\left(\frac{1}{L_C}+\frac{1}{L_D}\right)^2 + \left(\frac{d_\mathrm{D}}{L_D}\right)^2\right]\right) \tag{4}$$

Now the wavelength spread is described by $\Delta\lambda$ being the full width at the half maximum. The geometrical terms have contributions from the aperture sizes d_C and d_S and the spatial detector resolution d_D. The collimation length L_C and detector distance L_D are usually identical such that all geometric resolution contributions are evenly large ($d_C = 2d_S$ then). This ideal setup maximizes the intensity with respect to a desired resolution.
The resolution function profile is another topic of the correction calculations. A simple approach assumes Gaussian profiles for all contributions, and finally the overall relations read:

$$\left.\frac{d\Sigma(\bar{Q})}{d\Omega}\right|_{\mathrm{meas}} = \int_0^\infty dQ\; R(Q-\bar{Q}) \cdot \left.\frac{d\Sigma(Q)}{d\Omega}\right|_{\mathrm{theo}} \tag{5}$$

$$R(Q-\bar{Q}) = \frac{1}{\sqrt{2\pi}\sigma_Q}\exp\left(-\frac{1}{2}\frac{(Q-\bar{Q})^2}{\sigma_Q^2}\right) \tag{6}$$

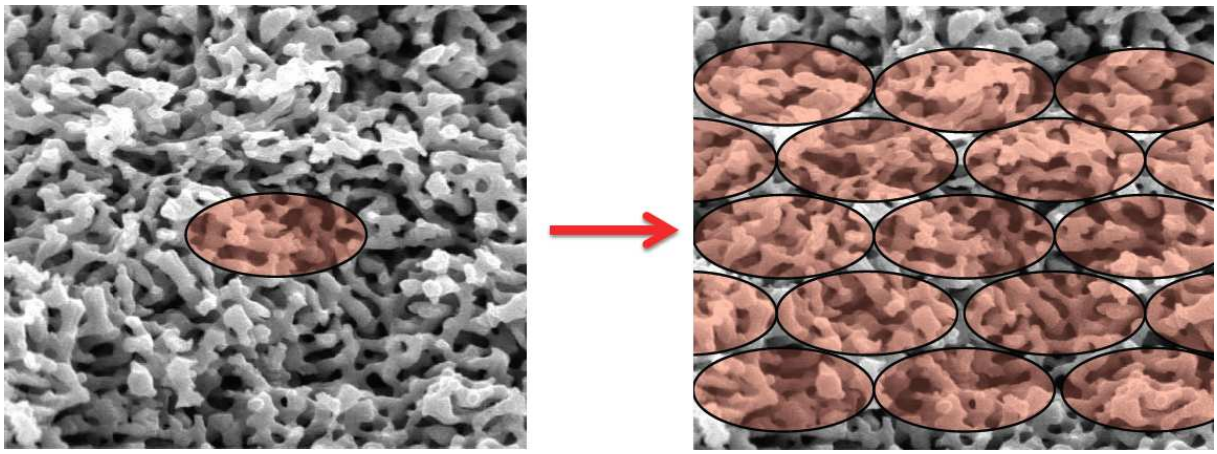


Fig. 7: *The coherence volume is usually much smaller than the sample volume (left). So the overall scattering appears as an incoherent superposition of the scattering from many coherence volumes (right).*

The theoretical macroscopic cross section is often described by a model function which is fitted to the experimental data. In this case the computer program only does a convolution of the model function with the resolution function $R(\Delta Q)$. Alternatively, there are methods to deconvolute the experimental data without modeling the scattering at first hand.
The here described resolution function is given as a Gaussian. This is true for relatively narrow distributions. The reason for using a Gaussian function although the original distributions of λ and θ are often triangular is: The central limit theorem can be applied to this problem because we have seen from eq. 4 that there are four contributions to the resolution function, and the radial averaging itself also smears the exact resolution function further out. Thus, the initial more detailed properties of the individual distributions do not matter anymore. Equations 5 and 6 are a good approximation for many practical cases.
We now want to describe the connection between the resolution function and the coherence of the neutron beam at the sample position. From optics we know about the transverse coherence length:

$$\ell_{\mathrm{coh,transv}} = \frac{\lambda L_C}{2 d_C} \qquad \text{is similar to} \qquad \Delta Q_\theta^{-1} = \frac{\lambda L_C}{\pi d_C} \tag{7}$$

It can be compared well with the geometric resolution contribution that arises from the entrance aperture only. Small differences in the prefactors we can safely neglect. For the longitudinal coherence length we obtain:

$$\ell_{\mathrm{coh,long}} = \frac{1}{4}\lambda \left(\frac{\Delta\lambda}{\lambda}\right)^{-1} \qquad \text{is similar to} \qquad \Delta k^{-1} = \frac{1}{2\pi}\lambda \left(\frac{\Delta\lambda}{\lambda}\right)^{-1} \tag{8}$$

This coherence length can be well compared to the wavevector uncertainty of the incoming beam. If we look back on Figure 4 we see that the coherence volume exactly describes the uncertainty of the incoming wave vector. The two contributions are perpendicular which supports the vectorial (independent) addition of the contributions in eq. 4 for instance. The coherence volume describes the size of the independent wave packages which allow for wave-like properties such as the scattering process. So the coherence volume describes the maximum size of

structure that is observable by SANS. If larger structures need to be detected the resolution must be increased.
The understanding how the small coherence volume covers the whole sample volume is given in the following (see also Fig. 7). Usually the coherence volume is rather small and is many times smaller than the irradiated sample volume. So many independent coherence volumes cover the whole sample. Then, the overall scattering intensity occurs as an independent sum from the scattering intensities of all coherence volumes. This is called incoherent superposition.

2.5 The theory of the macroscopic cross section (the Born approximation)

We have seen that the SANS instrument aims at the macroscopic cross section which is a function of the scattering vector $\mathbf{Q}$. In many examples of isotropic samples and orientationally averaged samples (powder samples) the macroscopic cross section depends on the modulus $|\mathbf{Q}| \equiv Q$ only. This measured function has to be connected to important structural parameters of the sample. For this purpose model functions are developed. The shape of the model function in comparison with the measurement already allows to distinguish the validity of the model. After extracting a few parameters with this method, deeper theories – like thermodynamics – allow to get deeper insight about the behavior of the sample. Usually, other parameters – like concentration, temperature, electric and magnetic fields, ... – are varied experimentally to verify the underlying concepts at hand. The purpose of this and the following sections is to give some ideas about model functions.
We have obtained a clear picture of the Born approximation in section 2.2. More formally, the Born approximation arises from quantum mechanics, and several facts and assumptions came along: The scattering amplitudes of the outgoing waves are derived as perturbations of the incoming plane wave. The matrix elements of the interaction potential with these two wave fields as vectors describe the desired amplitudes. The interaction potential can be simplified for neutrons and the nuclei of the sample by the Fermi pseudo potential. This expresses the smallness of the nuclei ($\sim$1fm) in comparison to the neutron wavelength ($\sim$Å). For the macroscopic cross section we immediately obtain a sum over all nuclei:

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left|\sum_j b_j \exp(\mathrm{i}\mathbf{Q}\cdot\mathbf{r}_j)\right|^2 \tag{9}$$

This expression is normalized to the sample volume V because the second factor usually is proportional to the sample size. This simply means: The more sample we put in the beam the more intensity we obtain. The second factor is the square of the amplitude because we measure intensities. While for electromagnetic fields at low frequencies one can distinguish amplitudes and phases (without relying on the intensity) the neutrons are quantum mechanical particles where experimentally such details are hardly accessible. For light (and neutrons) for instance holographic methods still remain. The single amplitude is a sum over each nucleus j with its typical scattering length b_j and a phase described by the exponential. The square of the scattering length b_j^2 describes a probability of a scattering event taking place for an isolated nucleus. The phase arises between different elementary scattering events of the nuclei for the large distances of the detector. In principle, the scattering length can be negative (for hydrogen for instance) which indicates an attractive interaction with a phase π. Complex scattering lengths indicate absorption. The quadrature of the amplitude can be reorganized:

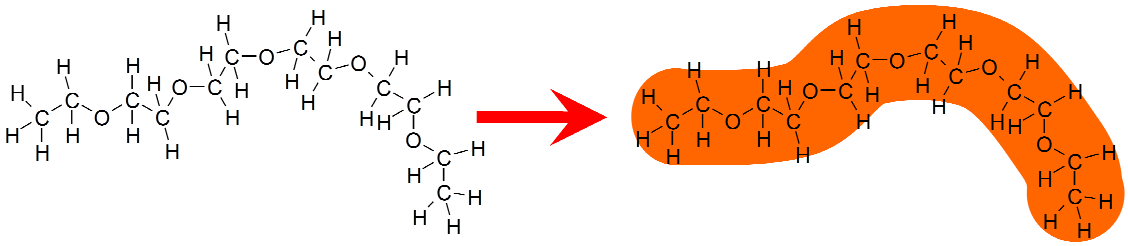


Fig. 8: *The concept of the scattering length density. On the left the atomic structure of a polyethylene oxide polymer (PEO) is depicted. For small angle scattering the wavelength is much larger than the atomic distance. So for SANS the polymer appears like a worm with a constant scattering length density inside.*

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\sum_{j,k} b_j b_k \exp\left(\mathrm{i}\mathbf{Q}(\mathbf{r}_j - \mathbf{r}_k)\right) \tag{10}$$

Here we find then self-terms with identical indices j and k without any phase and cross terms with phases arising from distances between different nuclei. Here it becomes obvious that only relative positions of the nuclei matter which is a result of the quadrature. The overall phase of the sample does not matter because of the modulus in eq. 9. We will use this expression for the polymer scattering.

Apart from this detailed expression a simplified view is allowed for small angle scattering experiments. Firstly, we know that the wavelength is typically 7Å which is much larger than the atom-atom distance of ca. 1.5Å. Secondly, the SANS experiment aims at structures at the nano-scale. So the scattering vector aims at much larger distances compared to the atomistic distances (i.e. $2\pi Q^{-1} \gg 1$Å). This allows for exchanging sums by integrals as follows:

$$\sum_j b_j \cdots \quad \longrightarrow \quad \int_V d^3r\ \rho(\mathbf{r}) \cdots \tag{11}$$

Such methods are already known for classical mechanics, but reappear all over physics. The meaning is explained by the sketch of Figure 8. The polymer polyethylene oxide (PEO) contains many different nuclei of different species (hydrogen, carbon and oxide). However, the SANS method does not distinguish the exact places of the nuclei. The polymer appears rather like a homogenous worm. Inside, the worm has a constant scattering length density which reads:

$$\rho_{\mathrm{mol}} = \frac{1}{V_{\mathrm{mol}}} \sum_{j\in\{\mathrm{mol}\}} b_j \tag{12}$$

So, for each molecule we consider all nuclei and normalize by the overall molecule volume. Of course different materials have different scattering length densities ρ. The initial equation 9 reads then:

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{1}{V}\left|\int_V d^3r\ \rho(\mathbf{r})\exp(\mathrm{i}\mathbf{Q}\mathbf{r})\right|^2 \tag{13}$$

$$= \frac{1}{V}\left|\mathcal{F}\left[\rho(\mathbf{r})\right]\right|^2 = \frac{1}{V}\left|\rho(\mathbf{Q})\right|^2 \tag{14}$$

The single amplitude is now interpreted as a Fourier transformation of the scattering length density $\rho(\mathbf{r})$ which we simply indicate by $\rho(\mathbf{Q})$. The amplitude is defined by:

$$\rho(\mathbf{Q}) = \int\limits_V d^3r\ \rho(\mathbf{r}) \exp(\mathrm{i}\mathbf{Q}\mathbf{r}) \tag{15}$$

Again, equation 13 loses the phase information due to the modulus.

2.6 Spherical colloidal particles

In this section we will derive the scattering of diluted spherical particles in a solvent. These particles are often called colloids, and can be of inorganic material while the solvent is either water or organic solvent. Later in the manuscript interactions will be taken into account.
One important property of Fourier transformations is that constant contributions will lead to sharp delta peaks at $Q = 0$. This contribution is not observable in the practical scattering experiment. The theoretically sharp delta peak might have a finite width which is connected to the overall sample size, but centimeter dimensions are much higher compared to the largest sizes observed by the scattering experiment ($\sim \mu$m). So formally we can elevate the scattering density level by any number $-\rho_{\mathrm{ref}}$:

$$\rho(\mathbf{r}) \longrightarrow \rho(\mathbf{r}) - \rho_{\mathrm{ref}} \qquad \text{leads to} \qquad \rho(\mathbf{Q}) \longrightarrow \rho(\mathbf{Q}) - 2\pi\rho_{\mathrm{ref}}\delta(\mathbf{Q}) \tag{16}$$

The resulting delta peaks can simply be neglected. For a spherical particle we then arrive at the simple scattering length density profile:

$$\rho_{\mathrm{single}}(\mathbf{r}) = \begin{cases} \Delta\rho & \text{for } |\mathbf{r}| \leq R \\ 0 & \text{for } |\mathbf{r}| > R \end{cases} \tag{17}$$

Inside the sphere the value is constant because we assume homogenous particles. The reference scattering length density is given by the solvent. This function will then be Fourier transformed accordingly:

$$\rho_{\mathrm{single}}(\mathbf{Q}) = \int\limits_0^{2\pi} d\phi \int\limits_0^{\pi} d\vartheta\ \sin\vartheta \int\limits_0^{R} dr\ r^2\ \Delta\rho\ \exp\left(\mathrm{i}|\mathbf{Q}| \cdot |\mathbf{r}| \cos(\vartheta)\right) \tag{18}$$

$$= 2\pi\ \Delta\rho \int\limits_0^{R} dr\ r^2 \left[\frac{1}{\mathrm{i}Qr} \exp\left(\mathrm{i}QrX\right)\right]_{X=-1}^{X=+1} \tag{19}$$

$$= 4\pi\ \Delta\rho \int\limits_0^{R} dr\ r^2\ \frac{\sin(Qr)}{Qr} \tag{20}$$

$$= \Delta\rho\ \frac{4\pi}{3} R^3 \left(3\frac{\sin(QR) - QR\cos(QR)}{(QR)^3}\right) \tag{21}$$

In the first line 18 we introduce spherical coordinates with the vector $\mathbf{Q}$ determining the z-axis for the real space. The vector product $\mathbf{Qr}$ then leads to the cosine term. In line 19 the

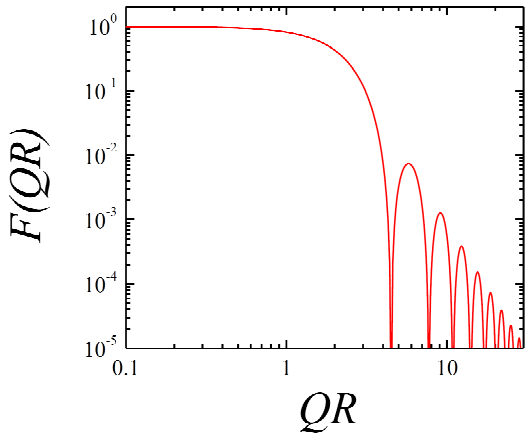


Fig. 9: *The form factor of a homogenous sphere in a double logarithmic plot.*

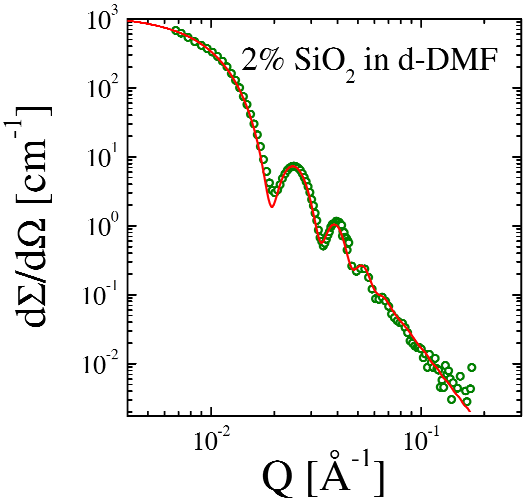


Fig. 10: *Experimental scattering curve of spherical SiO_2 colloids in the deuterated solvent DMF [4]. The resolution function (eq. 5) is included in the fit (red line).*

azimuthal integral is simply 2π, and the variable $X = \cos\vartheta$ is introduced. Finally, in line 20 the kernel integral for spherically symmetric scattering length density distributions is obtained. For homogenous spheres we obtain the final result of eq. 21. Putting this result together for the macroscopic cross section (eq. 14) we obtain:

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}) = \frac{N}{V} \cdot \left|\rho_{\text{single}}(\mathbf{Q})\right|^2 = (\Delta\rho)^2\, \phi_{\text{spheres}}\, V_{\text{sphere}}\, F(\mathbf{Q}) \tag{22}$$

$$F(\mathbf{Q}) = \left(3\frac{\sin(QR) - QR\cos(QR)}{(QR)^3}\right)^2 \tag{23}$$

We considered N independent spheres in our volume V, and thus obtained the concentration of spheres ϕ_{spheres}. Furthermore, we defined the form factor $F(Q)$, which describes the Q-dependent term for independent spheres (or the considered shapes in general). The function is shown in Figure 9. The first zero of the form factor is found at $Q = 4.493/R$. This relation again makes clear why the reciprocal space (Q-space) is called reciprocal. We know the limit for small scattering angles is $F(\mathbf{Q} \to 0) = 1 - \frac{1}{5}Q^2R^2$. So the form factor is normalized to 1, and the initial dependence on Q^2 indicates the size of the sphere. For large scattering angles the form factor is oscillating. Usually the instrument cannot resolve the quickest oscillations and an average intensity is observed. The asymptotic behavior would read $F(\mathbf{Q} \to \infty) = \frac{9}{2}(QR)^{-4}$. The obtained power law Q^{-4} is called Porod law and holds for any kind of bodies with sharp interfaces. So, sharp interfaces are interpreted as fractals with $d = 2$ dimensions, and the corresponding exponent is $6 - d$. The general appearance of the Porod formula reads then:

$$\frac{d\Sigma}{d\Omega}(Q) = P \cdot Q^{-4} \tag{24}$$

The amplitude of the Porod scattering P tells about the surface per volume and reads $P = 2\pi(\Delta\rho)^2 S_{\text{tot}}/V_{\text{tot}}$. Apart from the contrast, it measures the total surface S_{tot} per total volume V_{tot}. For our shperes, the Porod constant becomes $P = 2\pi(\Delta\rho)^2 4\pi R^2/(4\pi R^3/(3\phi)) =$

$6\pi\phi_1(\Delta\rho)^2/R$. The surface to volume ratio is smaller the larger the individual radius R is. The remaining scaling with the concentration ϕ_1 and the contrast $(\Delta\rho)^2$ arises still from the prefactor which we discussed in context with eq. 22.

When comparing the theoretical description of the spherical form factor with measurements one finds a good agreement (Fig. 10). Many fringes are seen, but after the third or fourth peak the function does not indicate any oscillation any more. Furthermore, the sharp minima are washed out. All of this is a consequence of the resolution function (eq. 5) which has been taken into account for the fitted curve. For many other examples one also needs to take the polydispersity into account. The synthesis of colloids usually produces a whole distribution of different radii. In our example the polydispersity is very low which is the desired case. Polydispersity acts in a similar way compared to the resolution function. The sharp minima are washed out. While the resolution appears as a distribution of different Q-values measured at a certain point the polydispersity integrates over several radii.

Another general scattering law for isolated (dilute) colloids is found for small scattering angles. The general appearance of the Guinier scattering law is:

$$\frac{d\Sigma}{d\Omega}(\mathbf{Q}\to 0) = \frac{d\Sigma}{d\Omega}(0)\cdot\exp\left(-\frac{1}{3}Q^2R_g^2\right) \tag{25}$$

When comparing the scattering law of a sphere and the Guinier formula we obtain $R_g = \sqrt{\frac{3}{5}}R$. The radius of gyration R_g can be interpreted as a momentum of inertia normalized to the total mass and specifies the typical size of the colloid of any shape. The Guinier formula can be seen as an expansion at small scattering angles of the logarithm of the macroscopic cross section truncated after the Q^2 term. Further details are discussed in Appendix A.

Another general appearance for independent colloids shall be discussed now using equation 22. The macroscopic cross section is determined by several important factors: The contrast between the colloid and the solvent given by $\Delta\rho^2$, the concentration of the colloids, the volume of a single colloid, and the form factor. Especially for small Q the latter factor turns to 1, and the first three factors dominate. When knowing two factors from chemical considerations, the third factor can be determined experimentally using small angle neutron scattering.

2.7 Scattering of a polymer

In this section we derive the scattering of a single (isolated) polymer coil. This model is the basis for many more complicated models of polymers in solution, polymeric micelles, polymer melts, diblock and multiblock copolymers and so on. So the understanding of these concepts is rather important for scattering experiments on any kind of polymer systems.

This example starts apart from many other calculations from point-like monomers (see eq. 10). These monomers are found along a random walk with an average step width of ℓ_K. We try to argue for non-ideal chain segments, but finally will arrive at an expression for rather ideal polymers. For the scattering function we obtain (definition of $S(Q)$ in eq. 37-39):

$$S(\mathbf{Q}) \propto \frac{1}{N}\sum_{j,k=0}^{N}\left\langle \exp\left(\mathrm{i}\mathbf{Q}\cdot(\mathbf{R}_j-\mathbf{R}_k)\right)\right\rangle \tag{26}$$

$$\propto \frac{1}{N}\sum_{j,k=0}^{N}\exp\left\langle -\tfrac{1}{2}\left(\mathbf{Q}\cdot(\mathbf{R}_j-\mathbf{R}_k)\right)^2\right\rangle \tag{27}$$

$$\propto \frac{1}{N}\sum_{j,k=0}^{N}\exp\left\langle -\tfrac{1}{6}\mathbf{Q}^2\cdot(\mathbf{R}_j-\mathbf{R}_k)^2\right\rangle \tag{28}$$

At this stage we use statistical arguments (i.e. statistical physics). The first rearrangement of terms (line 27) moves the ensemble average of the monomer positions (and distances $\Delta\mathbf{R}_{jk}$) from the outside of the exponential to the inside. This is an elementary step which is true for polymers. The underlying idea is, that the distance $\Delta\mathbf{R}_{jk}$ arises from a sum of $|j-k|$ bond vectors which all have the same statistics. So each sub-chain with the indices jk is only distinguished by its number of bond vectors inside. The single bond vector $\mathbf{b}_j$ has a statistical average of $\langle\mathbf{b}_j\rangle = 0$ because there is no preferred orientation. The next higher moment is the second moment $\langle\mathbf{b}_j^2\rangle = \ell_K^2$. This describes that each bond vector does a finite step with an average length of ℓ_K. For the sub-chain we then find an average size $\langle\Delta\mathbf{R}_{jk}^2\rangle = |j-k|\ell_K^2$. The reason is that in the quadrature of the sub-chain only the diagonal terms contribute because two distinct bond vectors show no (or weak) correlations.
Back to the ensemble average: The original exponential can be seen as a Taylor expansion with all powers of the argument $\mathrm{i}\mathbf{Q}\Delta\mathbf{R}_{jk}$. The odd powers do not contribute with similar arguments than for the single bond vector $\langle\mathbf{b}_j\rangle = 0$. Thus, the quadratic term is the leading term. The reason why the higher order terms can be arranged that they finally fit to the exponential expression given in line 27 is the weak correlations of two distinct bond vectors. The next line 28 basically expresses the orientational average of the sub-chain vector $\Delta\mathbf{R}_{jk}$ with respect to the Q-vector in three dimensions.
This derivation can be even simpler understood on the basis of a Gaussian chain. Then every bond vector follows a Gaussian distribution (with a center of zero bond length). Then the ensemble average has the concrete meaning $\langle\cdots\rangle = \int\cdots\exp\left(-\frac{3}{2}\Delta\mathbf{R}_{jk}^2/(|j-k|\ell_K^2)\right)d^3\Delta R_{jk}$. This distribution immediately explains the rearrangement of line 27. The principal argument is the central limit theorem: When embracing several segments as an effective segment any kind of distribution converges to yield a Gaussian distribution. This idea came from Kuhn who formed the term Kuhn segment. While elementary bonds still may have correlations at the stage of the Kuhn segment all correlations are lost, and the chain really behaves ideal. This is the reason why the Kuhn segment length ℓ_K was already used in the above equations.
In the following we now use the average length of sub-chains (be it Kuhn segments or not), and replace the sums by integrals which is a good approximation for long chains with a large number of segments N.

$$S(\mathbf{Q}) \propto \frac{1}{N}\int_0^N dj\int_0^N dk\exp\left(-\tfrac{1}{6}\mathbf{Q}^2\cdot|j-k|\cdot\ell_K^2\right) \tag{29}$$

$$= N\cdot f_D(Q^2R_g^2) \tag{30}$$

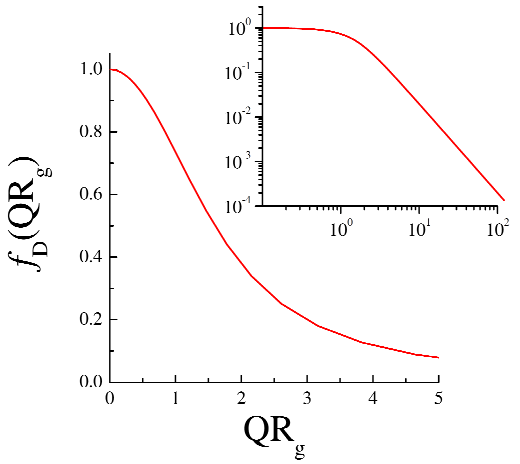


Fig. 11: *The theoretical Debye function describes the polymer scattering of independent polymers without interaction. The two plots show the function on a linear and double logarithmic scale.*

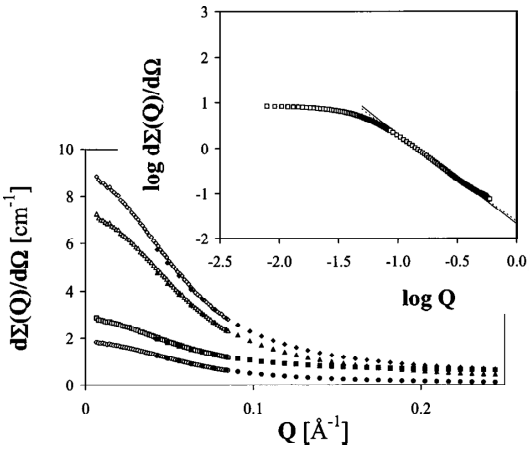


Fig. 12: *Scattering of a h/d-PDMS polymer blend. The linear scale shows different compositions of hydrogenous polymer (from bottom to top: 0.05, 0.94, 0.27, 0.65) while the double logarithmic plot shows the 0.65 sample only [5].*

$$f_D(x) = \tfrac{2}{x^2}\left(\exp(-x) - 1 + x\right) \qquad \text{with} \quad x = Q^2 R_g^2 \tag{31}$$

In this integral one has to consider the symmetry of the modulus. The result is basically the Debye function which describes the polymer scattering well from length scales of the overall coil down to length scales where the polymer becomes locally rigid (see Fig. 11). The covalent bonds of a carbon chain effectively contribute to a certain rigidity which will not be treated here. The radius of gyration describes the overall dimension of the chain and is $R_g = \sqrt{N/6}\,\ell_K$. The limits of the polymer scattering are found to be:

$$S(\mathbf{Q}) \propto N(1 - \tfrac{1}{3}Q^2R_g^2) \qquad \text{for small } Q \tag{32}$$

$$\propto N \cdot 2/(Q^2R_g^2) \qquad \text{for large } Q \tag{33}$$

The line 32 describes the conventional Guinier scattering of the overall polymer (compare eq. 25). The second line 33 describes a power law. At these length scales the sub-chains of different lengths are self-similar and so they reveal a fractal behavior. The prefactor is connected to the magnitude R_g^2/N which is the effective segment size. From this magnitude one can calculate back to the local rigidity which is responsible for the effective segments.
When we want to compare experiments with this theory the best examples are obtained from polymer blends (Fig. 12). One could come to the conclusion that diluted polymer solutions must provide the ideal conditions for such an experiment but practically the interactions of the solvent molecules with the monomers lead to a deviating behavior: The good solvent conditions lead to energetic violations of monomer-monomer contacts and so the polymer swells and displays a different fractal behavior. The high Q power law in good solvents comes close to $Q^{-1.7}$. The Flory theory was the first attempt to describe this behavior while many refinements find small corrections. The theoretically most precise Flory exponent is $\nu = 0.588$ which is the reciprocal value of the given exponent 1.7 above.

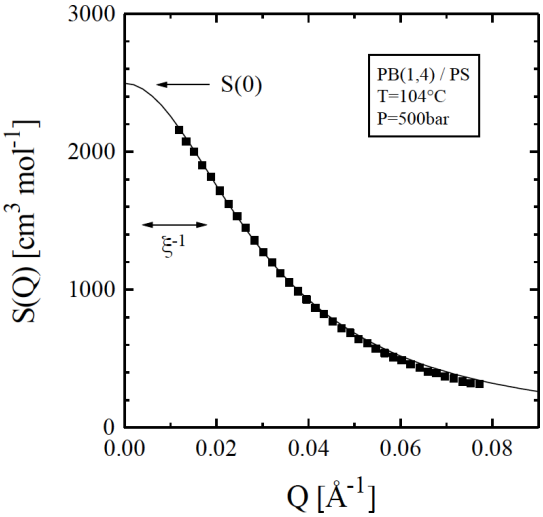


Fig. 13: *Typical scattering of a homopolymer blend with interactions. The sample is a polybutadiene(1,4) / polystyrene blend at 104°C and 500bar [6].*

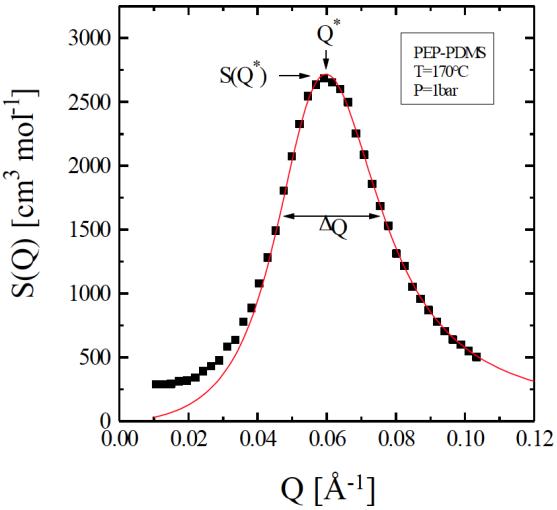


Fig. 14: *Typical scattering of a diblock copolymer blend with interactions [7]. The poly-ethylene-propylene–poly-dymethylsiloxane is heated to 170°C (1bar).*

So polymer blends are often better examples for weakly interacting chains. This finding is supported by the low entropy of mixing which enforces small interactions. The discussed example of Fig. 12 [5] considers the isotopic mixture of hydrogenous and deuterated polydimethylsiloxane (PDMS). This practically leads to one of the lowest possible interactions even though they are not completely zero. The theoretical concept of the random phase approximation is able to deal with interactions and describes phase diagrams and the scattering in this way. At high temperatures the polymers usually mix well, and the scattering comes closer to the weakly interacting case. Closer to the demixing temperature at lower temperatures the scattering intensity increases dramatically. This indicates strong composition fluctuations. The system loses the tendency to form a homogenous mixture and so local enrichments of species A or B are possible. While the random phase approximation is a mean field concept which describes weak fluctuations there are other concepts for strong fluctuations close to the phase boundary: The 3-dimensional Ising model – known for ferromagnets – describes the strong fluctuations of the two component polymer system.

The example of an interacting homopolymer blend is shown in Fig. 13. The general aspects are kept from non-interacting polymers (compare Fig. 12). The scattering curve has a maximum at $Q = 0$, and is decaying to large Q where a power law of Q^{-2} for ideal chains is observed. The maximal intensity is connected to the reciprocal susceptibility which describes the tendency of spontaneous thermal fluctuations to decay. High intensities mean low susceptibilies and strong fluctuations – the vice versa arguments are valid. The width of this curve is connected to the correlation length ξ. At low interactions it is tightly connected to the single coil size, i.e. R_g. With strong fluctuations close to the phase boundary the correlation length tends to diverge, which measures the typical sizes of the thermally fluctuating enrichments.

A diblock copolymer is a linear chain with two different monomer species. The first part is pure A and the latter pure B. The typical scattering of a diblock copolymer blend is shown in Fig. 14. At small Q the ideal scattering increases with Q^2 accordingly to the ‘correlation holeł. The chemistry of the molecule does not allow for enrichments of A or B on large length scales. A continuously growing volume would only allow for enrichments on the surface – this

explains finally the exponent in the scattering law. The experimental finite intensities relate to imperfections of the molecules. The chain length ratio f is distributed, and finally allows for enrichments on large length scales.
The dominating fluctuations are found at a finite Q^*. This expresses that the coil allows for separations of A and B predominantly on the length scale of the overall coil. Close to the phase boundary and especially below (where the polymer undergoes a micro phase separation) the coils are stretched. The peak at finite Q^* also expresses that the fluctuations tend to form alternating enrichments. From a center it would look like a decaying order of A-B-A-B-... The width of the peak is again connected to the correlation length $\xi \sim \Delta Q^{-1}$ which describes the length of this decaying order. At high Q, again a Q^{-2} law is observed describing the sub-chains being ideal chains. On these length scales homopolymer blends and diblock copolymers do not differ.
The whole understanding of phase boundaries and fluctuations is important for applications. Many daily life plastic products consist of polymer blends since the final product should have combined properties of two different polymers. Therefore, polymer granulates are mixed at high temperatures under shear in an extruder. The final shape is given by a metal form where the polymer also cools down. This process involves a certain temperature history which covers the one-phase and two-phase regions. Therefore, the final product consists of many domains with almost pure polymers. The domain size and shape are very important for the final product. So the process has to be tailored in the right way to yield the specified domain structure. This tailoring is supported by a detailed knowledge of the phase boundaries and fluctuation behavior. Advanced polymer products also combine homopolymers and diblock copolymers for an even more precise and reproducible domain size/shape tailoring [8, 9]. The diblock copolymer is mainly placed at the domain interfaces, and, therefore, influences the domain properties precisely.

2.8 Contrast variation

For neutron scattering the method contrast variation opens a wide field of possible experiments. For soft matter research the most important labeling approach is the exchange of hydrogen ^{1}H by deuterium ^{2}H $\equiv$ D. Since in a single experiment the phase information is lost completely the contrast variation experiment retrieves this information partially. Relative positions of two components are obtained by this method.
The scattering length density of the overall sample is now understood to originate from each component individually. So the specific $\rho_j(\mathbf{r})$ takes the value of the scattering length density of component j when the location points to component j and is zero otherwise. We would then obtain the following:

$$\rho(\mathbf{Q}) = \int_V d^3r \left(\sum_{j=1}^{n} \rho_j(\mathbf{r}) \right) \exp(\mathrm{i}\mathbf{Q}\mathbf{r}) \tag{34}$$

n specifies the number of components. The assumption of incompressibility means that on every place there is one component present, and so all individual functions $\rho_j(\mathbf{r})$ fill the full space. Furthermore, we would like to define component 1 being the reference component, i.e. $\rho_{\mathrm{ref}} = \rho_1$ (see eq. 16). This means that on each place we have a $\Delta\rho_j(\mathbf{r})$ function similar to eq. 17. Then, we arrive at:

$$\rho(\mathbf{Q}) = \sum_{j=2}^{n} \Delta\rho_{j1}(\mathbf{Q}) \tag{35}$$

The macroscopic cross section is a quadrature of the scattering length density $\rho(\mathbf{Q})$, and so we arrive at:

$$\begin{aligned} \frac{d\Sigma}{d\Omega}(\mathbf{Q}) &= \frac{1}{V} \cdot \sum_{j,k=2}^{n} \Delta\rho_{j1}^{*}(\mathbf{Q}) \cdot \Delta\rho_{k1}(\mathbf{Q}) && (36)\\ &= \sum_{j,k=2}^{n} (\Delta\rho_{j1}\Delta\rho_{k1}) \cdot S_{jk}(\mathbf{Q}) && (37)\\ &= \sum_{j=2}^{n} (\Delta\rho_{j1})^2 \cdot S_{jj}(\mathbf{Q}) \; + \; 2 \sum_{2<j<k\leq n} (\Delta\rho_{j1}\Delta\rho_{k1}) \cdot \Re\, S_{jk}(\mathbf{Q}) && (38) \end{aligned}$$

In line 37 the scattering function $S_{jk}(\mathbf{Q})$ is defined. By this the contrasts are separated from the Q-dependent scattering functions. Finally, in line 38 the diagonal and off-diagonal terms are collected. There are $n-1$ diagonal terms, and $\frac{1}{2}(n-1)(n-2)$ off-diagonal terms. Formally, these $\frac{1}{2}n(n-1)$ considerably different terms are rearranged (the combinations $\{j,k\}$ are now simply numbered by j), and a number of s different measurements with different contrasts are considered.

$$\left.\frac{d\Sigma}{d\Omega}(\mathbf{Q})\right|_s = \sum_{j} (\Delta\rho \cdot \Delta\rho)_{sj} \cdot S_j(\mathbf{Q}) \tag{39}$$

In order to reduce the noise of the result, the number of measurements s exceeds the number of independent scattering functions considerably. The system then becomes over-determined when solving for the scattering functions. Formally, one can nonetheless write:

$$S_j(\mathbf{Q}) = \sum_{s} (\Delta\rho \cdot \Delta\rho)^{-1}_{sj} \cdot \left.\frac{d\Sigma}{d\Omega}(\mathbf{Q})\right|_s \tag{40}$$

The formal inverse matrix $(\Delta\rho \cdot \Delta\rho)^{-1}_{sj}$ is obtained by the singular value decomposition method. It describes the closest solution of the experiments in context of the finally determined scattering functions.

An example case is discussed for a bicontinuous microemulsion with an amphiphilic polymer [10]. The microemulsion consists of oil and water domains which have a sponge structure. So the water domains host the oil and vice versa. The surfactant film covers the surface between the oil and water domains. The symmetric amphiphilic polymer position and function was not clear beforehand. From phase diagram measurements it was observed that the polymer increases the efficiency of the surfactant dramatically. Much less surfactant is needed to solubilize equal amounts of oil and water. Fig. 15 discusses the meaning of the cross terms of the scattering functions. Especially the film-polymer scattering is highly interesting to reveal the polymer role inside the microemulsion (see Fig. 16). By the modeling it was clearly observed that the amphiphilic polymer is anchored in the membrane and the two blocks describe a mushroom inside the oil and water domains. So basically, the polymer is a macro-surfactant. The coils of the polymer exert a certain pressure on the membrane and keep it flat. The membrane with less

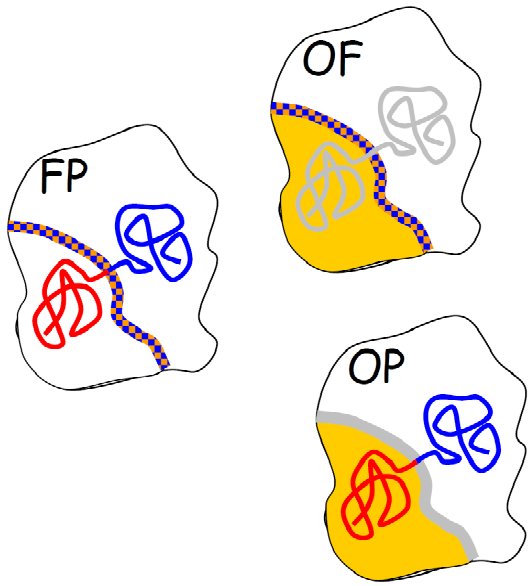


Fig. 15: *Scheme of scattering functions for the cross terms within the microemulsion. There are the film-polymer scattering S_{FP}, the oil-film scattering S_{OF}, and the oil-polymer scattering S_{OP}. The real space correlation function means a convolution of two structures.*

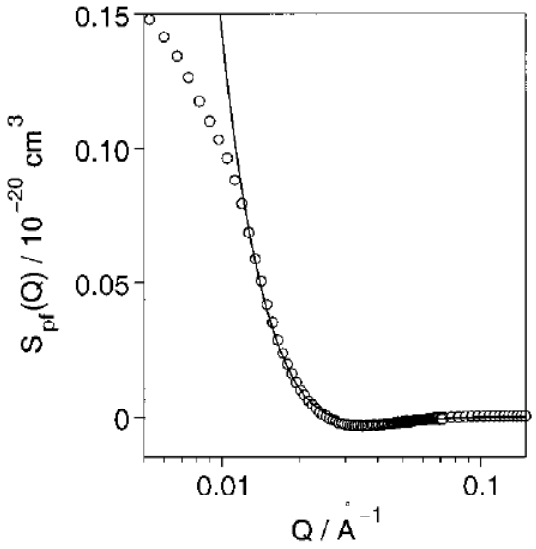


Fig. 16: *A measurement of the film-polymer scattering for a bicontinuous microemulsion with a symmetric amphiphilic polymer. The solid line is described by a polymer anchored in the film. The two blocks are mushroom-like in the domains. At low Q the overall domain structure (or size) limits the idealized model picture.*

fluctuations allows for the formation of larger domains with a better surface to volume ratio. This is finally the explanation how the polymer acts as an efficiency booster.

3 Small Angle X-ray Scattering

While a detailed comparison between SANS and SAXS is given below, the most important properties of the small angle x-ray scattering technique shall be discussed here. The x-ray sources can be x-ray tubes (invented by Röntgen, keyword Bremsstrahlung) and modern synchrotrons. The latter ones guide fast electrons on undulators which act as laser-like sources for x-rays with fixed wavelength, high brilliance and low divergence. This simply means, that the collimation of the beam often yields narrow beams, and the irradiated sample areas are considerably smaller (often smaller than ca. $1\times1\mathrm{mm}^2$). A view on the sample position is given in Fig. 17 (compare Fig. 3). One directly has the impression that all windows are tiny and adjustments must be made more carefully.

The conceptual understanding of the scattering theory still holds for SAXS. For the simplest understanding of the contrast conditions in a SAXS experiment, it is sufficient to count the electron numbers for each atom. The resulting scattering length density reads then (compare eq. 12):

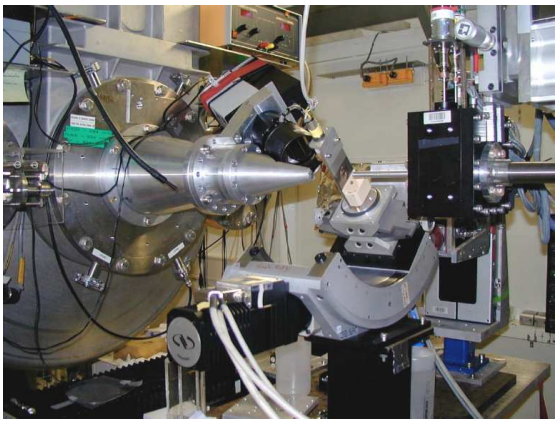


Fig. 17: *The sample position of the SAXS instrument ID2 at the ESRF, Grenoble, France. The photons propagate from the right to the left. The collimation guides on the left and the detector tank window on top of the cone on the left give an impression about the small beam size (being typically $1\times1\text{mm}^2$).*

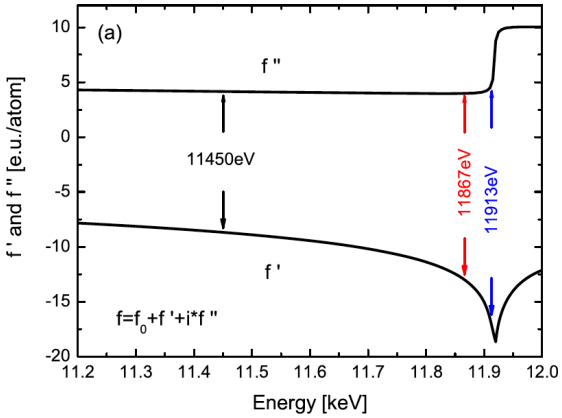


Fig. 18: *The complex dispersion curve for gold (Au) at the L_3 edge [11]. The overall effective electron number $f = f_0 + f' + \mathrm{i}f''$ replaces the conventional electron number $Z = f_0$ in equation 41. On the x-axis the energies of the x-rays is shown, with indications for the experimentally selected three energies (black, red, blue). In this way, equal steps for the contrast variation are achieved.*

$$\rho_{\text{mol}} = \frac{r_e}{V_{\text{mol}}} \sum_{j\in\{\text{mol}\}} Z_j \tag{41}$$

The classical electron radius is $r_e = \mathrm{e}^2/(4\pi\epsilon_0 m_e c^2) = 2.82\text{fm}$. The electron number of each atom j is Z_j. This means that chemically different substances have a contrast, but for similar substances (often for organic materials) it can be rather weak. Heavier atoms against light materials are much easier to detect. Finally, the density of similar materials is also important. Especially for organic materials (soft matter research), the high intensity of the source still allows for collecting scattering data. Many experiments base on these simple modifications with respect to SANS, and so the fundamental understanding of SAXS experiments does not need any further explanation.

For completeness, we briefly discuss the scattering length density for light scattering. Here the polarizability plays an important role. Without going into details, the final contrast is expressed by the refractive index increment dn/dc:

$$\rho_{\text{mol}} = \frac{2\pi n}{\lambda^2} \cdot \frac{dn}{dc_{\text{mol}}} \tag{42}$$

The refractive index increment dn/dc finally has to be determined separately experimentally when the absolute intensity is of interest. The concentration c_{mol} is given in units volume per volume (for the specific substance in the solvent). The wavelength of the used light is λ.

3.1 Contrast variation using anomalous small angle x-ray scattering

While for contrast variation SANS experiments the simple exchange of hydrogen ^{1}H by deuterium $^2\mathrm{H} \equiv \mathrm{D}$ allowed for changing the contrast without modifying the chemical behavior, in contrast variation SAXS experiments the applied trick is considerably different: The chemistry is mainly dominated by the electron or proton number Z and isotope exchange would not make any difference. The electron shells on the other hand have resonances with considerable dispersion curves. An example is shown in Fig. 18 with the real part f' (called dispersion) and the imaginary part f'' (called absorption). The overall effective electron number $f = f_0 + f' + \mathrm{i}f''$ replaces the conventional electron number $Z = f_0$ in equation 41. Below the resonance energy the considered L_3 shell appears only softer and effectively less electrons appear for f. Above the resonance energy single electrons can be scattered out from the host atom (Compton effect). This is directly seen in the sudden change of the absorption. Furthermore, the actual dependence of the dispersion is influenced by backscattering of the free electrons to the host atom (not shown in Fig. 41). This effect finally is the reason that the complex dispersion curve can only theoretically be well approximated below the resonance (or really far above). For this approximation it is sufficient to consider isolated host atoms.
For best experimental results the f-values have to be equally distributed. Thus, the energies are selected narrower close to the resonance (see Fig. 41). The investigated sample consisted of core-shell gold-silver nanoparticles in soda-lime silicate glass (details in reference [11]). By the contrast variation measurement one wanted to see the whole particles in the glass matrix, but also the core-shell structure of the individual particles. Especially, the latter one would be obtained from such an experiment. First results of this experiment are shown in Fig. 19. The most important result from this experiment is that the original scattering curves at first hand do not differ considerably. The core-shell structure results from tiny differences of the measurements. For contrast variation SANS experiments the contrasts can be selected close to zero contrast for most of the components which means that tiniest amounts of additives can be highlighted and the intensities between different contrasts may vary by factors of 100 to 1000. So for contrast variation SAXS measurements the statistics have to be considerably better which in turn comes with the higher intensities.
Another example was evaluated to a deeper stage [12]. Here, the polyelectrolyte polyacrylate (PA) with Sr^{2+} counterions was dissolved in water. The idea behind was that the polymer is dissolved well in the solvent. The charges of the polymer and the ions lead to a certain swelling of the coil (exact fractal dimensions ν not discussed here). The counterions form a certain cloud around the chain – the structure of which is the final aim of the investigation. The principles of contrast variation measurements leads to the following equation (compare eq. 39):

$$\frac{d\Sigma}{d\Omega} = (\Delta\rho_{\mathrm{Sr-H2O}})^2 \cdot S_{\mathrm{Sr-Sr}} \; + \; (\Delta\rho_{\mathrm{PA-H2O}})^2 \cdot S_{\mathrm{PA-PA}} \; + \; \Delta\rho_{\mathrm{Sr-H2O}}\Delta\rho_{\mathrm{PA-H2O}} \cdot S_{\mathrm{Sr-PA}} \quad (43)$$

The overall scattering is compared with two contributions in Fig. 20. The scattering functions of the cross term $S_{\mathrm{Sr-PA}}$ and the pure ion scattering $S_{\mathrm{Sr-Sr}}$ have been compared on the same scale, and so the contrasts are included in Fig. 20. Basically, all three functions describe a polymer coil in solvent – the different contrasts do not show fundamental differences. Nonetheless, a particular feature of the ion scattering was highlighted by this experiment: At $Q \approx 0.11\mathrm{nm}^{-1}$ is a small maximum which is connected to the interpretation of effective charge beads along the chains. The charge clouds obviously can be divided into separated beads. The emphasis of the observed maximum correlates with the number of beads: For small numbers it is invisible,

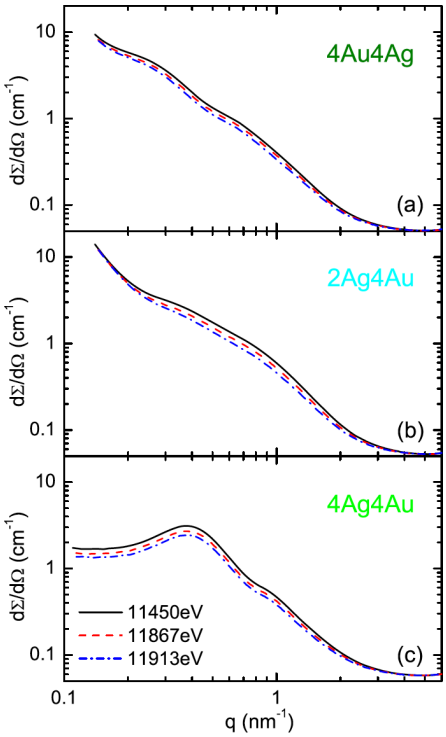


Fig. 19: *Absolute calibrated scattering curves of different core-shell Ag/Au nanoparticles in soda-lime silicate glass [11]. The implantation sequence has been changed for the three samples. Note that the three scattering curves for the selected energies (colors correspond to Fig. 18) do only slightly differ due to the small changes of the contrast.*

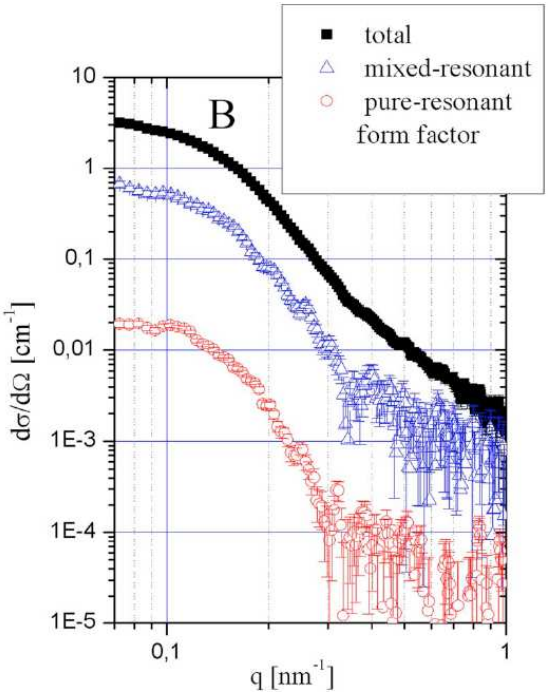


Fig. 20: *Further evaluated scattering functions of a different system [12]: A polyelectrolyte with Sr^{2+} counterions in aqueous solution. The top curve (black) indicates the overall scattering. The middle curve (blue) displays the polymer-ion cross terms being sensitive for relative positions. The bottom curve (red) depicts the pure ion scattering.*

and becomes more pronounced with higher numbers. The authors finally find that the number of 5 beads is suitable for the description of the scattering curves: An upper limit is also given by the high Q scattering where the 5 chain segments appear as independent sub-coils. This example beautifully displays, that the method of contrast variation can be transferred to SAXS experiments. Difficulties of small contrast changes have been overcome by the good statistics due to much higher intensities.

4 Comparison of SANS and SAXS

We have seen that many parallels exist between the two experimental methods SANS and SAXS. The theoretical concepts are the same. Even the contrast variation method as a highly difficult and tedious task could be applied for both probes. In the following, we will highlight differences that have been discussed so far and others that are just mentioned now.

The high flux reactors are at the technical limit of highest neutron fluxes. For SANS instruments maximal fluxes of ca. 2×10^8 neutrons/s/cm^2 have been reached at the sample position. Typical sample sizes are of 1×1cm^2. For coherent scattering fractions of ca. 10% this results in maximal

count rates of 10^7Hz, while practically most of the count rates stay below 10^6Hz. For long collimations, the experimentalists deal often with 10 to 50Hz. The resolution for these count rates has been relaxed. Wavelength spreads of either $\pm$5% or $\pm$10% are widely accepted, and the collimation contributes equally, such that a typical resolution of $\Delta Q/Q$ of 7 to 14% is reached. For many soft matter applications this is more than adequate. If one thinks of liquid crystalline order, much higher resolution would be desired which one would like to overcome by choppers in combination with time-of-flight analysis. A resolution of ca. 1% would be a reasonable expectation. The continuous sources are highly stable which is desired for a reliable absolute calibration.

The spallation sources deliver either continuous beams or the most advanced ones aim at pulsed beams. Repetition rates range from ca. 14 to 60Hz. The intensity that is usable for SANS instruments could reach up to 20 times higher yields (as planned for the ESS in Lund), i.e. up to 4×10^9 neutrons/s/cm^2. Surely, detectors for count rates of 10 to 100 MHz have to be developed. The new SANS instruments will make use of the time-of-flight technique for resolving the different wavelengths to a high degree. Of course other problems with such a broad wavelength band have to be overcome – but this topic would lead too far.

The synchrotron sources reach much higher photon yields which often makes the experiments technically comfortable but for the scientist at work highly stressful. The undulators provide laser-like qualities of the radiation which explains many favorable properties. Some numbers for the SAXS beam line ID2 at the ESRF shall be reported. The usable flux of 5×10^{15} photons/s/mm^2 (note the smaller area) is provided which results for a typical sample area of ca. 1×0.02mm^2 in 10^{14} photons/s. In some respect the smallness of the beam urges to think about the representativeness of a single shot experiment. At some synchrotron sources the beam is not highly stable which makes absolute calibration and background subtraction difficult. The same problem also occurs for the pulsed neutron sources where parts of the calibration procedure become highly difficult.

For classical SANS experiments one can make some statements: The absolute calibration is practically done for all experiments and does not take much effort – it is technically simple. Between different instruments in the world the discrepancies of different calibrations results often in errors of 10% and less. Part of the differences are different calibration standards, but also different concepts for transmission measurements and many details of the technical realization. The nuclear scattering is a result of the fm small nuclei and results in easily interpretable scattering data for even large angles – for point-like scatterers no corrections have to be made. In this way all soft matter and biological researchers avoid difficult corrections. Magnetic structures can be explored by neutrons due to its magnetic moment. Magnetic scattering is about to be implemented to a few SANS instruments. Ideally, four channels are experimentally measured (I_{++}, I_{+-},I_{-+}, and I_{--}) by varying the polarization of the incident beam (up/down) and of the analyzer. Nowadays, the ^{3}He technique allows for covering relatively large exit angles at high polarization efficiencies. But also early magnetic studies have been possible with simpler setups and reduced information. The unsystematic dependence of the scattering length often opens good conditions for a reasonable contrast for many experiments. If the natural isotopes do not provide enough contrast pure isotopes might overcome the problem. The contrast variation experiments have been presented for the SANS technique. By a simple exchange of hydrogen by deuterium, soft matter samples can be prepared for complicated contrast variation experiments. One advantage is the accessibility of the zero contrast for most of the components which allows for highlighting smallest amounts of additives. The high demand for deuterated

chemicals makes them cheap caused by the huge number of NMR scientists. The low absorption of neutrons for many materials allows for studying reasonably thick samples (1 to 5mm and beyond). Especially, for contrast variation experiments often larger optical path lengths are preferred. The choice for window materials and sample containers is simple in many cases. Neutron scattering is a non-destructive method. Espeically biological samples can be recovered.

Contrarily we observe for the SAXS technique: The demand for absolute calibration in SAXS experiments is growing. Initial technical problems are overcome and suitable calibration standards have been found. The interpretation of scattering data at larger angles might be more complicated due to the structure of the electron shells. For small angle scattering the possible corrections are often negligible. Magnetic structures are observable by the circular magnetic dichroism [13] but do not count to the standard problems addressed by SAXS. The high contrast of heavy atoms often makes light atoms invisible. For soft matter samples the balanced use of light atoms results in low contrast but, technically, the brilliant sources overcome any intensity problem. The ASAXS technique is done close to resonances of single electron shells and opens the opportunity for contrast variation measurements. The achieved small differences in the contrast still allow for tedious measurements because the statistics are often extremely good – only stable experimental conditions have to be provided. The absorption of x-rays makes the choice of sample containers and windows more complicated. The absorbed radiation destroys the sample in principle. Short experimental times are thus favorable.

To summarize, the method of small angle neutron scattering is good-natured and allows to tackle many difficult tasks. The small angle x-ray scattering technique is more often applied due to the availability. Many problems have been solved (or will be solved) and will turn to standard techniques. So, in many cases the competition between the methods is kept high for the future. Today, practically, the methods are complementary and support each other for the complete structural analysis.

Appendices

A Guinier Scattering

The crucial calculation of the Guinier scattering is done by a Taylor expansion of the logarithm of the macroscopic cross section for small scattering vectors Q. Due to symmetry considerations there are no linear terms, and the dominating term of the Q-dependence is calculated to be:

$$
\begin{aligned}
R_g^2 &= -\frac{1}{2}\cdot\frac{\partial^2}{\partial Q^2}\ln\Big(\rho(\mathbf{Q})\rho(-\mathbf{Q})\Big)\Big|_{\mathbf{Q}=0} && (44)\\
&= -\frac{1}{2}\cdot\frac{\partial}{\partial Q}\frac{2\Re\big(\rho(\mathbf{Q})\int d^3r\ \rho(\mathbf{r})(-\mathrm{i}\mathbf{r})\exp(-\mathrm{i}\mathbf{Q}\mathbf{r})\big)}{\rho(\mathbf{Q})\rho(-\mathbf{Q})}\Big|_{\mathbf{Q}=0} && (45)\\
&= -\Re\frac{\rho(\mathbf{Q})\int d^3r\ \rho(\mathbf{r})(-\mathbf{r}^2)\exp(-\mathrm{i}\mathbf{Q}\mathbf{r})}{\rho(\mathbf{Q})\rho(-\mathbf{Q})}\Big|_{\mathbf{Q}=0}\\
&\quad -\Re\frac{\int d^3r\ \rho(\mathbf{r})(\mathrm{i}\mathbf{r})\exp(\mathrm{i}\mathbf{Q}\mathbf{r})\int d^3r\ \rho(\mathbf{r})(-\mathrm{i}\mathbf{r})\exp(-\mathrm{i}\mathbf{Q}\mathbf{r})}{\rho(\mathbf{Q})\rho(-\mathbf{Q})}\Big|_{\mathbf{Q}=0} + 0 && (46)\\
&= \langle\mathbf{r}^2\rangle - \langle\mathbf{r}\rangle^2 && (47)\\
&= \Big\langle\big(\mathbf{r}-\langle\mathbf{r}\rangle\big)^2\Big\rangle && (48)
\end{aligned}
$$

The first line 44 contains the definition of the Taylor coefficient. Then, the derivatives are calculated consequently. Finally, we arrive at terms containing the first and second momenta. The last line 48 rearranges the momenta in the sense of a variance. So the radius of gyration is the second moment of the scattering length density distribution with the center of 'gravity' being at the origin. We used the momenta in the following sense:

$$
\begin{aligned}
\langle\mathbf{r}\rangle &= \int d^3r\ \mathbf{r}\rho(\mathbf{r})\Big/\int d^3r\ \rho(\mathbf{r}) && (49)\\
\langle\mathbf{r}^2\rangle &= \int d^3r\ \mathbf{r}^2\rho(\mathbf{r})\Big/\int d^3r\ \rho(\mathbf{r}) && (50)
\end{aligned}
$$

So far we assumed an isotropic scattering length density distribution. In general, for oriented anisotropic particles, the Guinier scattering law would read:

$$
\frac{d\Sigma}{d\Omega}(\mathbf{Q}\to 0) = \frac{d\Sigma}{d\Omega}(0)\cdot\exp\Big(-Q_x^2\Big\langle\big(x-\langle x\rangle\big)^2\Big\rangle - Q_y^2\Big\langle\big(y-\langle y\rangle\big)^2\Big\rangle - Q_z^2\Big\langle\big(z-\langle z\rangle\big)^2\Big\rangle\Big) \quad (51)
$$

Here, we assumed a diagonal tensor of second moment. This expression allows for different widths of scattering patterns for the different directions. In reciprocal space large dimensions appear small and vice versa. Furthermore, we see that R_g is defined as the sum over all second momenta, and so in the isotropic case a factor $\frac{1}{3}$ appears in the original formula 25.

References

[1] R.-J. Roe, *Methods of X-ray and Neutron Scattering in Polymer Science* (Oxford University Press, New York, 2000).

[2] J.S. Higgins, H.C. Benoît, *Polymers and Neutron Scattering* (Clarendon Press, Oxford, 1994).

[3] J. Skov Pedersen, D. Posselt, K. Mortensen, *J. Appl. Cryst.* **23**, 321 (1990).

[4] T. Vad, W.F.C. Sager, J. Zhang, J. Buitenhuis, A. Radulescu, *J. Appl. Cryst.* **43**, 686 (2010).

[5] S. Gagliardi, V. Arrighi, R. Ferguson, A.C. Dagger, J.A. Semlyen, J.S. Higgins, *J. Chem. Phys.* **122**, 064904 (2005).

[6] H. Frielinghaus, D. Schwahn, L. Willner, *Macromolecules* **34**, 1751 (2001).

[7] H. Frielinghaus, D. Schwahn, K. Mortensen, K. Almdal, T. Springer, *Macromolecules* **29**, 3263 (1996).

[8] J.H. Lee, M.L. Ruegg, N.P.Balsara, Y. Zhu, S.P. Gido, R. Krishnamoorti, M.H. Kim *Macromolecules* **36**, 6537 (2003) and B.J. Reynolds, M.L. Ruegg, N.P. Balsara, C.J. Radke, T.D. Shaffer, M.Y. Lin, K.R. Shull, D.J. Lohse *Macromolecules* **37**, 7410 (2004)

[9] N.R. Washburn, T.P. Lodge, F.S. Bates *J. Phys. Chem. B* **104**, 6987 (2000)

[10] H. Endo, M. Mihailescu, M. Monkenbusch, J. Allgaier, G. Gompper, D. Richter, B. Jakobs, T. Sottmann, R. Strey, I. Grillo, *J. Chem. Phys.* **115**, 580 (2001).

[11] J. Haug, H. Kruth, M. Dubiel, H. Hofmeister, S. Haas, D. Tatchev, A. Hoell, *Nanotechnology* **20**, 505705 (2009).

[12] G. Goerigk, K. Huber, R. Schweins, *J. Chem. Phys.* **127**, 154908 (2007) and G. Goerigk, R. Schweins, K. Huber, M. Ballauff, *Europhys. Lett.* **66**, 331 (2004).

[13] P. Fischer, R. Zeller, G. Schütz, G. Goerigk, H.G. Haubold, *J. Phys. IV France* **7**, C2-753 (1997).

D 2 Scattering under Grazing Incidence from Surfaces and Interfaces

U. Rücker
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Motivation

During the last decades, the possibility to create structures confined along one or more dimensions to a size in the nanometer range has induced a class of materials with extraordinary properties, as well from the scientific as from the technological points of view. Beginning from Si based integrated circuits over layered magnetic structures showing the Giant Magnetoresistance (GMR) effect [1] to displays out of conducting polymers produced by ink-jet technology [2], layered structures use the interface effects on the conducting electrons and (in the case of magnetic structures) on the electron spin.

For a complete understanding of the properties of a layered structure the approximation of laterally homogeneous layers with sharp boundaries generally does not hold. For example in the case of the magnetic exchange coupling between a ferromagnetic and an antiferromagnetic layer, which leads to a shift of the hysteresis loop of the ferromagnetic layer ("exchange bias"), the measured effect is two orders of magnitude weaker than the calculated effect based on the assumption of a magnetic structure in the antiferromagnet which is not affected by the interface. The up-to-date models need to take into account monoatomic steps at the interface between ferromagnet and antiferromagnet and formation of magnetic domains inside the antiferromagnet [3,4,5].

Like this, there are many examples of effects in thin film structures, e.g. chemical segregation, magnetic interlayer coupling, tunnelling magnetoresistance, proximity effects, spin injection, etc., where it is important to be able to investigate the chemical and magnetic structure in the inner layers of the sample. Most of the methods used to investigate thin film structures either integrate over the sample (e.g. SQUID magnetometry), are purely surface sensitive (e.g. Scanning Tunneling Microscopy STM, Atomic Force Microscopy AFM, Magnetic Force Microsopy MFM, Low Energy Electron Diffraction LEED) or integrate over a certain depth (e.g. Auger Electron Spectroscopy AES, Kerr microscopy). Some methods are element specific (e.g. Photoemission Electron Microscopy PEEM) without pronounced depth sensitivity, so they allow to distinguish between different layers consisting of different elements, but not to distinguish between several layers with identical composition, e.g. between the two Fe layers in a Fe / Cr / Fe trilayer.

For a specific access to the buried interfaces or layers, we need to use a probe that can penetrate the sample and that is coherent over the size of the objects of interest to be able to detect the interference between signals from different depth. Useful probes are x-ray photons as well as neutrons, because they can penetrate solid samples and have wavelengths in the order of a few Å, suitable to resolve structures of atomic size.

Generally, in thin film research, the atomic structure of the layers is not subject to the investigation, but the layer structure and inhomogeneties of the layers, e.g. growth islands, density fluctuations, magnetic domains or interdiffusion regions.

To resolve layer structure, vertical interface profiles and interdiffusion regions, one needs a resolution in the z-direction of the order of a fraction of the layer thickness. As layer thicknesses vary from atomic monolayers to several 100 nm, a z-resolution in the order of Å to µm is needed.

In the lateral direction, i.e. parallel to the film plane, different structures are of interest. It starts with lithographically produced structures in the µm range, grains in polycrystalline layers (1000 to 50 nm), growth islands (100 nm to 10 nm) and goes down to self-organized lateral structures (down to 1 nm). Sometimes, e.g. in self-organized nanoparticle arrays deposited on a substrate, several length scales are of importance. The internal structure of the array (i.e. the periodicity of the nanoparticle arrangement) is in the range of several nanometers, while the size distribution of the domains of coherent structure or the clusters of nanoparticles can extend up to several micrometers. All these length ranges are accessible with scattering under grazing incidence.

2 Specular reflectivity and scattering under grazing incidence

The investigation of structure and interfaces of thin films with x-rays or neutrons is mostly performed using elastic scattering under grazing incidence. In this case, a monochromatic, well collimated beam impinges under a well defined, small angle $\alpha_i = \theta$ (in most cases $\theta << 5°$) onto the surface of the sample. It is then partly reflected specularly from the surface, i.e. the outgoing angle $\alpha_f = \theta$ as well, and partly refracted into the material (see Fig. 1). As I will derive in section 4, the reflection of x-rays or neutrons from a laterally homogeneous medium can be treated according to classical optics. Only the proper index of refraction n for the radiation has to be used.

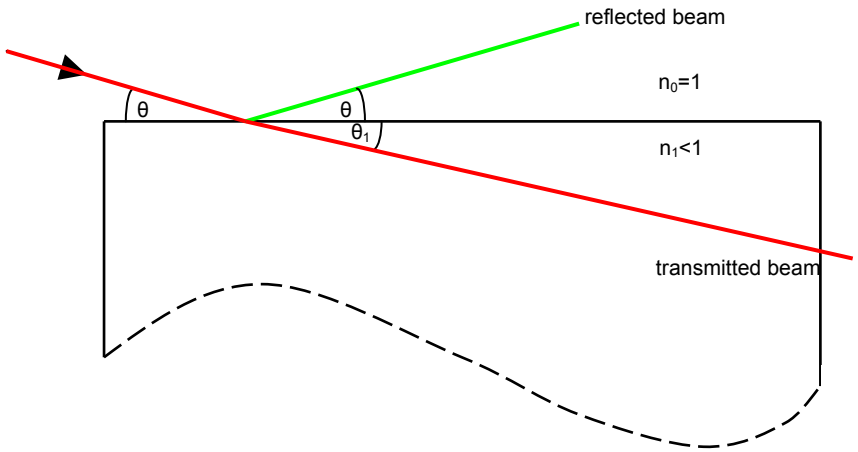


Fig. 1: *Reflection and Refraction from a free surface*

For most materials, the index of refraction for neutrons and x-rays is slightly smaller than 1, leading to total external reflection for small angles of incidence $\theta < \theta_c$, where θ_c depends on the material and the radiation.

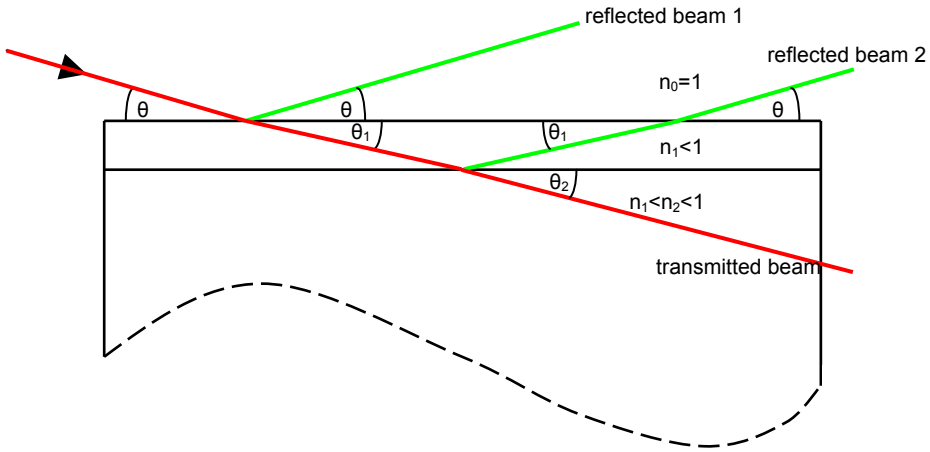


Fig. 2: *Reflection and Refraction from a single layer on a substrate*

In the case of a single layer on the substrate, reflection and refraction take place at both the surface and the interface (Fig. 2). Then, the reflected beams from the different interfaces interfere with each other. Maximum intensity is received, when the path length difference between the two reflected beams is an integer multiple of the wavelength. Due to the continuity relation of the wave function at the interfaces, the intensity of the transmitted beam is modulated opposite to the reflected beam. Fig. 10 shows the reflectivity curve of a real single layer on a substrate. It is discussed in detail in chapter 5.1.

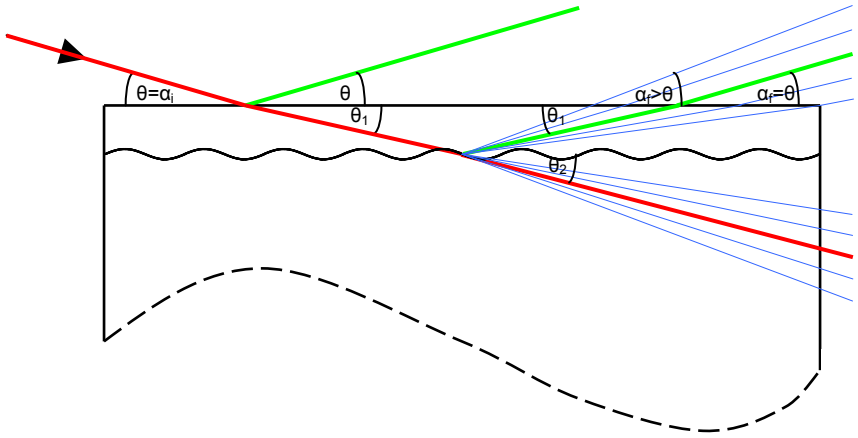


Fig. 3: *Specular and off-specular scattering from a laterally modulated interface*

If surface and interface are laterally ideally homogeneous and flat, the complete beam is either reflected specularly or transmitted after refraction. If the interface is not flat, but e.g. periodical (as sketched in Fig. 3), additional beams are coming up. Their origin is diffraction from the modulated interface, as it is known from an optical grating. This so-called off-specular or diffuse scattering can be observed at $\alpha_f \neq \alpha_i$.

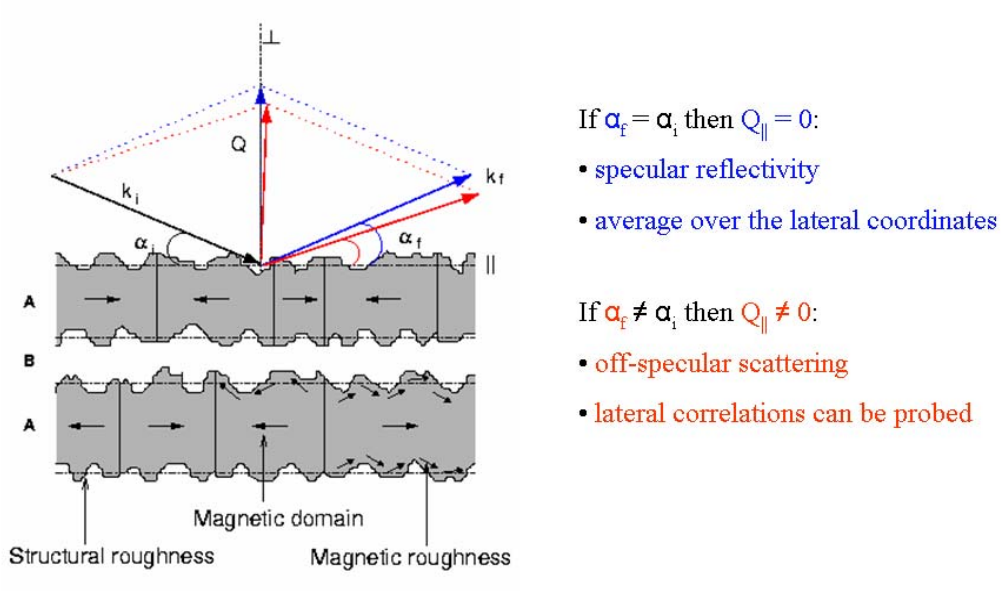


Fig. 4: *General geometry of reflectivity and off-specular scattering*

The general principle of a scattering experiment under grazing incidence is depicted in Fig. 4 and the exact geometry is given in Fig. 5. For the mathematical description of the scattering process, it is again convenient to introduce the wave vectors $\mathbf{k_i}$ and $\mathbf{k_f}$ for the incoming and the detected wave, respectively. Neutrons or x-rays impinge on the sample surface under the grazing angle of incidence α_i with wave vector $\mathbf{k_i}$ and are detected with outgoing wave vector $\mathbf{k_f}$ whose direction is defined by the angle α_f in the plane of incidence and by the angle φ perpendicular to it. Let $\mathbf{Q} = \mathbf{k_f} - \mathbf{k_i}$ be the scattering wave vector and $\mathbf{Q}_{//} = (Q_x,Q_y,0)$ its component in the sample plane. I will restrict the discussion to elastic scattering, i.e. $k_i = k_f = 2\pi / \lambda$ where λ is the wavelength of the radiation.

If the sample can be considered as laterally homogeneous, i.e. invariant by translation along its surface, intensity can only be observed in the specular direction defined by $\mathbf{Q}_{//} = 0$, i.e. at $\alpha_f = \alpha_i$ and $\varphi = 0$. If the sample shows lateral fluctuations like chemical roughness, magnetic roughness or magnetic domains, then some intensity can be observed in the directions given by $\mathbf{Q}_{//} \neq 0$, i.e. by $\alpha_f \neq \alpha_i$ and $\varphi \neq 0$. Very often specular reflectivity and scattering under grazing incidence are observed simultaneously. Specular reflectivity then gives information

on the order parameters averaged over the lateral coordinates and diffuse scattering gives access to the fluctuations around this mean value.

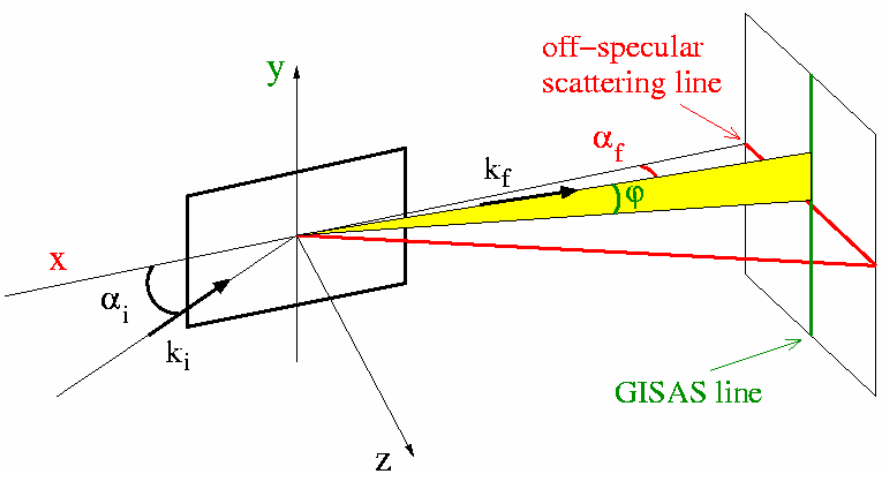


Fig. 5: *The different geometries of scattering under grazing incidence. Off-specular scattering probes lateral correlations along the x-direction, while grazing incidence small angle scattering (GISAS) probes correlations along the x- and y-directions.*

In practice, two types of scattering under grazing incidence geometries can be used (Fig. 5). The first one is obtained by scanning α_i and α_f while integrating the measured intensities along φ. This is called off-specular scattering. For the second one, the experimental conditions are such that φ can be resolved. Recording the intensities along φ as a function of α_i and α_f is called grazing incidence small angle scattering (GISAS).

Taking into account the small values of the angles α_i, α_f and φ, the scattering wave vector projects itself on the three axis of the coordinate system of Fig. 5 in the following manner:

$$\mathbf{Q} = \begin{cases} Q_x \approx k \cdot (\alpha_i^2 - \alpha_f^2 - \varphi^2) \\ Q_y \approx k \cdot \varphi \\ Q_z \approx k \cdot (\alpha_i + \alpha_f) \end{cases} \qquad (1)$$

Note that while Q_x is a linear combination of squares of small angles, Q_y is linear in φ. Q_x is then always much smaller than Q_y. Typically, at a neutron source one has 0.5 µm < $d_x = 2\pi/Q_x$ < 20 µm and 1 nm < $d_y = 2\pi/Q_y$ < 300 nm. The lower limits are defined by the maximum available intensity and the upper limits are fixed by the reachable resolution in Q_x and Q_y. It appears then that GISANS (grazing incidence small angle neutron scattering) probes much smaller length scales than off-specular scattering. For x-rays the lower limits for d_x and d_y can

be brought even smaller thanks to the availability of high intensity synchrotron sources. As GISAS gives access to the nanometer length scale in the sample plane, one observes at the moment a rush on this method, with both neutrons (GISANS) and x-rays (GISAXS).

Off-specular and GISAS intensities are represented differently, because they have a different dimension. The 3-dimensional off-specular data are typically represented as a function of α_i and α_f with a colour encoding of the intensity (Fig. 6a). Along the main diagonal where $\alpha_i = \alpha_f$ is the specular line ($Q_x = 0$). Out of this diagonal, off-specular scattering is measured and the lateral correlations are probed along the x-direction ($Q_x \neq 0$). Along lines perpendicular to the specular line where $\alpha_i + \alpha_f$ = constant, the correlations are probed along Q_x at Q_z constant (cf. eq. (1)).

GISAS measurements are 4-dimensional and therefore cannot be displayed in a single figure containing all data. As the experiments are usually performed on a small angle scattering instrument [6] where the data is collected on the 2D position sensitive detector, it is a common way to present the measurements at fixed α_i. For a well defined angle of incidence α_i and a well defined sample-to-detector distance, the intensities recorded on the 2D detector give the GISAS signal as a function of $\alpha_i + \alpha_f$ and φ (Fig. 6b). As $Q_y \approx k \cdot \varphi$, the φ axis directly gives access to the correlations along the y-axis. The specular reflectivity peak can be found at the coordinate ($\alpha_i + \alpha_f = 2\,\alpha_i$, $\varphi = 0$).

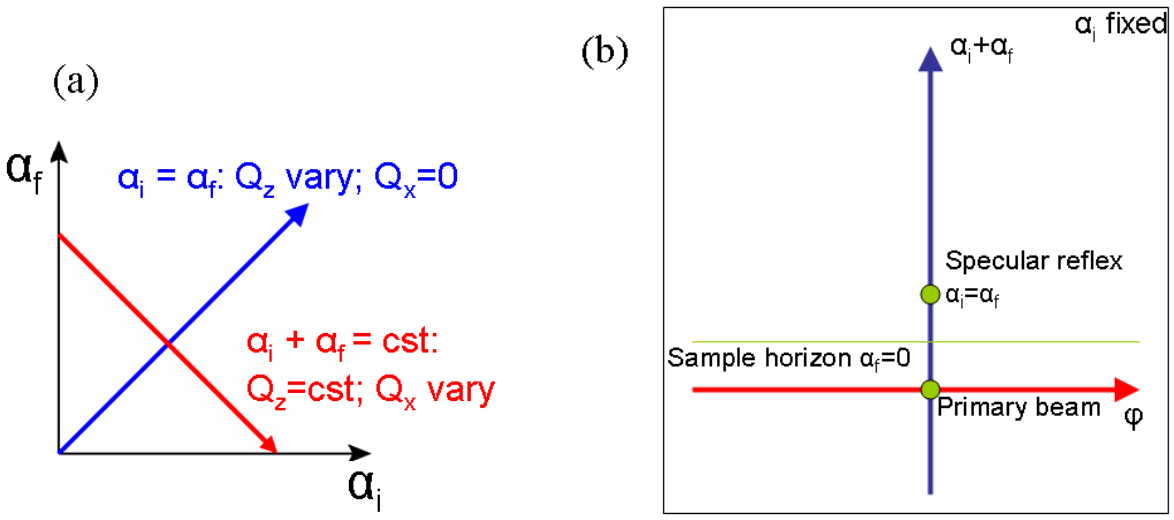


Fig. 6: *Data representation of off-specular scattering (a) and GISAS (b)*

3 Spatial coherence of the beam

An important quantity to consider in the interpretation of specular reflectivity and scattering under grazing incidence is the coherence volume of the beam [7], because its projection S_{coh} on the sample surface, while sometimes large enough to induce multiple scattering effects, stays smaller than the illuminated part of the sample surface. This is always true for a neutron beam. At shallow angles it might not be true for a x-ray beam at a synchrotron source, because of the high spatial coherence of the x-ray beam due to the large distance of the source with respect to the sample, its small lateral size and the good wavelength resolution. In that case, the surface of coherence is the entire illuminated part of the sample surface.

$$S_{coh} = L^{x}_{coh} \cdot L^{y}_{coh} \tag{2}$$

is connected to the resolutions of the in-plane components of the scattering wave vector

$$L^{x}_{coh} \propto 1/\delta Q_x \quad \text{and} \quad L^{y}_{coh} \propto 1/\delta Q_y . \tag{3}$$

which themselves depend on the quality of monochromatization (wavelength spread) of the beam and the angular divergence. These values always depend on the instrument used and the settings of the slits etc. that should be optimized for the specific experiment.

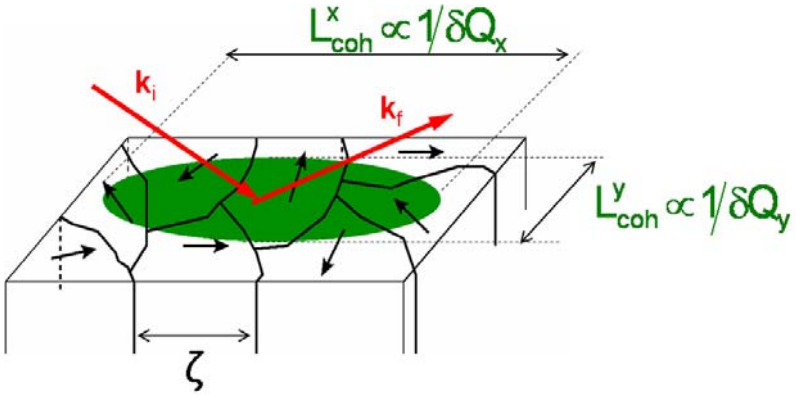


Fig. 7: *Projection of the coherence volume of the beam on the sample surface. The surface thus obtained has the lateral dimensions* L^{x}_{coh} *and* L^{y}_{coh} *along the x- and y-directions of Fig. 5. In order to interpret correctly the measured data, those lengths have to be compared with the correlation length* ζ *of the lateral fluctuations.*

If, within S_{coh}, the sample consists of homogeneous layers with flat interfaces, there cannot be any transfer between the in-plane components of $\mathbf{k}_i$ and $\mathbf{k}_f$ and intensity is observed only along the specular direction defined by $\mathbf{Q}_{//} = \mathbf{0}$ with a precision given by $\delta\mathbf{Q}_{//}$. The measured reflectivity is therefore an average over all the reflectivities generated by all the surfaces of coherence that compose the surface of the sample and for which the homogeneous and flat approximation holds. Therefore, the absence of scattering under grazing incidence is not always due to the fact that the sample is homogeneous and flat over its whole surface.

If, on the contrary, the system is not invariant by a lateral translation along the x or y direction by a vector whose modulus is smaller than L^x_{coh} or L^y_{coh}, respectively, a transfer of momentum parallel to the surface can take place ($\mathbf{Q}_{//} \neq 0$) and scattering under grazing incidence can be observed. The scattering cross section is then an average over the scattering cross sections generated by the different surfaces of coherence that compose the sample.

Usually, scattering under grazing incidence and specular reflectivity coexist. Reflectivity comes from the depth variation of the order parameter averaged over the lateral coordinates and scattering under grazing incidence finds its origin in the fluctuations around this mean value.

4 Specular reflectivity: Optical approach

For the case of a perfectly smooth surface, an exact description of the reflected and transmitted intensity can be deduced from quantum theory. This approach is also valid, if the correlation lengths of the fluctuations exceed largely the coherence length in the respective direction.

As an example, I will show the case for neutrons, although the identical calculus is valid for x-rays. Only some nature constants are different, resulting in a different term describing the index of refraction.

The starting point is the Schrödinger equation for the wave function of the neutron

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{4}$$

The energy of the neutron is given by $E = \hbar^2 k^2/(2m)$ with the modulus $k = 2\pi/\lambda$ of the wave vector $\mathbf{k}$. As we assume elastic scattering, the energy of the incident and of the outgoing wave is identical.

The resolution of a reflectometry experiment does not resolve the atomic structure of the sample in any of the three directions. Therefore, it is a valid approximation to describe the

potential V_1 of the homogeneous material as the sum of the scattering length densities of all constituents (see lecture A4):

$$V_1 = \frac{2\pi\hbar^2}{m}\sum_j b_j\rho_j \tag{5}$$

where b_j are the coherent scattering lengths and ρ_j the atomic number densities of the different elements (evtl. isotopes) in the material. With that, we receive

$$\left[\Delta + \left(k^2 - 4\pi\sum_j b_j\rho_j\right)\right]\Psi(\mathbf{r}) = \left[\Delta + k^2\left(1 - \frac{\lambda^2}{\pi}\sum_j b_j\rho_j\right)\right]\Psi(\mathbf{r}) = \left[\Delta + k_1^2\right]\Psi(\mathbf{r}) = 0 \tag{6}$$

with the wave vector $\mathbf{k}_1$ inside the medium. From this equation, it is justified to introduce the index of refraction of the material

$$n = \frac{k_1}{k} \qquad n = 1 - \frac{\lambda^2}{2\pi}\sum_j b_j\rho_j =: 1 - \delta \qquad \delta = \frac{\lambda^2}{2\pi}\sum_j b_j\rho_j \tag{7}$$

It is a number very close to 1 for thermal and cold neutrons. The correction δ is called dispersion and is in the order of 10^{-5} to 10^{-6}. For most materials δ is positive (because the coherent scattering length b is positive for most isotopes), so that n is smaller than 1. This means that the transmitted beam is refracted towards the sample surface, which is opposite to the daily experience with light refracted at a glass or liquid surface.

For most materials, such as silicon, aluminium or iron, the absorption of neutrons is negligible. In case, it is not negligible, it can be introduced most straightforward by including an imaginary part to the index of refraction:

$$n = 1 - \delta + i\beta \tag{8}$$

In the case of x-rays the description in the framework of optical refraction inside the material is valid as well. In this case the index of refraction can be calculated as

$$n = 1 - \frac{\lambda^2 r_o}{2\pi}\sum_j \rho_j(Z_j + f_j' + if_j'') =: 1 - \delta + i\beta \qquad \delta = \frac{\lambda^2 r_0}{2\pi}\sum_j \rho_j(Z_j + f_j') \tag{9}$$

r_0 is the classical electron radius $r_0 = e/m_e c^2 = 2.82$ fm, Z is the number of electrons of the atom and f' and f'' are corrections for dispersion and absorption close to resonance energies. Typically, they can be neglected, only at the absorption edges they become important. An example for contrast variation close to the resonances is given in section 5.3.
Also for x-rays, the dispersion δ is always positive, so that the index of refraction n is smaller than 1.

In analogy to classical optics, we can derive e.g. Fresnel's formulas: For the solution of the wave equation at the sharp interface, we assume the surface of the sample to be at z = 0. The potential is then

$$V(z) = \begin{cases} 0 & \text{for } z > 0 \\ V_1 & \text{for } z \le 0 \end{cases} \tag{10}$$

As the potential is independent on x and y, the wave vector $\mathbf{k}_l$ in the wave equation (6) is also independent on x and y. Therefore, the wave equation can be separated by the Ansatz

$$\Psi(\mathbf{r}) = \Psi_x(x)\Psi_y(y)\Psi_z(z) \tag{11}$$

For the z direction we receive the one-dimensional differential equation

$$\Psi_z''(z) + k_z^2(z)\Psi_z(z) = 0 \tag{12}$$

To solve the differential equation we use the Ansatz

$$\Psi_{zl}(z) = t_l e^{ik_{zl}z} + r_l e^{-ik_{zl}z} \tag{13}$$

The index l distinguishes between vacuum (l=0) and matter (l=1). The factors t_l describe the wave field away from the surface, i.e. the transmitted wave, the factors r_l describe the wave field towards the surface, i.e. the reflected wave. The unique solution is determined by the boundary conditions. In a half-infinite medium, there is no reflected wave, because there is nothing to reflect from, i.e. r_l vanishes. In addition, the wave function and its first derivative must be continuous at the interface. So we receive the boundary conditions

$$\begin{aligned} &\Psi_{z0}(z=0) = \Psi_{z1}(z=0) \\ &\Psi_{z0}'(z=0) = \Psi_{z1}'(z=0) \\ &r_1 = 0 \end{aligned} \tag{14}$$

When we insert (14) into (12) and (13), we receive the continuity equations for the wave function

$$\begin{aligned} &t_0 + r_0 = t_1 \\ &k_{z0}(t_0 - r_0) = k_{z1}t_1 \end{aligned} \tag{15}$$

t_0 is the amplitude of the incoming wave, t_1 of the transmitted wave and r_0 of the reflected wave.

We can rewrite this set of equations in a matrix equation

$$\begin{pmatrix} t_0 \\ r_0 \end{pmatrix} = \mathbf{M_0} \begin{pmatrix} t_1 \\ r_1 \end{pmatrix} = \mathbf{M_0} \begin{pmatrix} t_1 \\ 0 \end{pmatrix} \tag{16}$$

The reflectivity R is defined as the ratio of the intensities of reflected and incoming waves, the transmissivity T is defined as the ratio of the intensities of transmitted and incoming waves.

$$R = \left| \frac{r_0}{t_0} \right|^2 \qquad T = \left| \frac{t_1}{t_0} \right|^2 \tag{17}$$

In conclusion, we arrive at Fresnel's formulas for the reflection at a flat interface

$$\text{Reflectivity} \qquad R = \left| \frac{k_{z0} - k_{z1}}{k_{z0} + k_{z1}} \right|^2 = \left| \frac{\theta - n_1 \theta_1}{\theta + n_1 \theta_1} \right|^2 \tag{18}$$

$$\text{Transmissivity} \qquad T = \left| \frac{2k_{z0}}{k_{z0} + k_{z1}} \right|^2 = \left| \frac{2\theta}{\theta + n_1 \theta_1} \right|^2 \tag{19}$$

Taking into account the continuity relation for the wave vector component tangential to the surface

$$k_{x0} = k_{x1} \qquad k_{y0} = k_{y1} \tag{20}$$

together with $k_1 = k_0 n$ (eq. 7), Snell's law for refraction follows from trigonometry:

$$\frac{\cos\theta_0}{\cos\theta_1} = \frac{k_1}{k_0} = n_1 \tag{21}$$

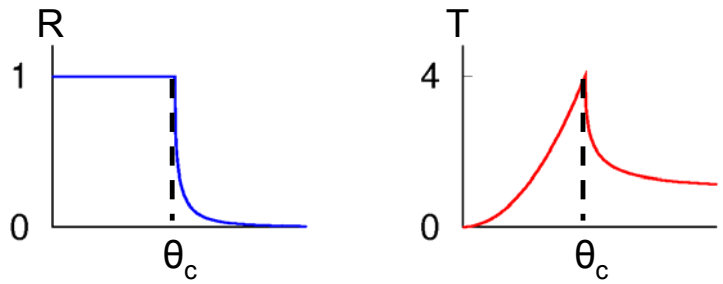


Fig. 8: *Reflectivity and transmittivity as a function of the angle of incidence*

The fact that in most cases the index of refraction is $n < 1$ means that the transmitted beam is refracted towards the sample surface ($\theta_1 < \theta$ in Fig. 1). For angles of incidence θ below the so-called critical angle θ_c with

$$n = \cos\theta_c \qquad \theta_c \approx \sqrt{2\delta} \tag{22}$$

total reflection is observed, i.e. all intensity is reflected and no wave propagating in z-direction exists in the sample. Only an evanescent wave with propagation parallel to the surface is induced. In this case, t_1 as well as k_{z1} are imaginary numbers. Still, as shown in Fig. 8, the modulus of t_1 is increasing when approaching θ_c. The strong evanescent wave field inside the surface is the origin of the strong offspecular Yoneda scattering close to $\alpha_i = \theta_c$ and $\alpha_f = \theta_c$. For incident angles above θ_c, the beam can partially penetrate the sample and is only partly reflected.

In the case of p layers on a substrate, the same calculus can be used. At every interface, the continuity relation can be formulated analogous to (16):

$$\begin{pmatrix} t_0 \\ r_0 \end{pmatrix} = \mathbf{M_0}\begin{pmatrix} t_1 \\ r_1 \end{pmatrix} \qquad \begin{pmatrix} t_1 \\ r_1 \end{pmatrix} = \mathbf{M}_1\begin{pmatrix} t_2 \\ r_2 \end{pmatrix} \qquad \ldots\ldots \qquad \begin{pmatrix} t_p \\ t_p \end{pmatrix} = \mathbf{M}_p\begin{pmatrix} t_{p+1} \\ 0 \end{pmatrix} \tag{23}$$

The amplitudes of reflected and transmitted wave then can be calculated by a matrix multiplication of the individual reflection matrices:

$$\begin{pmatrix} t_0 \\ r_0 \end{pmatrix} = \mathbf{M_0} \cdot \mathbf{M_1} \cdot \ldots \cdot \mathbf{M_p}\begin{pmatrix} t_{p+1} \\ 0 \end{pmatrix} \tag{24}$$

From equation (18), it becomes obvious, that reflection is only achieved when a difference between the indices of refraction of the adjacent layers exists. If one wants to investigate two layers with neutron or x-ray reflectivity, the contrast of the indices of refraction is the important quantity deciding, if the interface is visible or not. The higher the contrast, i.e. the higher the difference between the two indices of refraction, the higher is the contribution of this interface to the reflectivity curve.

The contrast achievable decides, which probe is useful to investigate a certain structure. Fig. 9 gives an impression of the scattering lengths of different elements for neutrons and x-rays. The scattering length density, which is proportional to the dispersion δ as well for neutrons as for x-rays, is proportional to the density and to the scattering length of the elements contained in the respective layer. For x-rays, the scattering length is generally proportional to the number of electrons, while for neutrons it is quite randomly distributed over the periodic system and over the different isotopes of each element. Light atoms as well as neighbouring atoms in the periodic system cannot be distinguished well with x-rays.

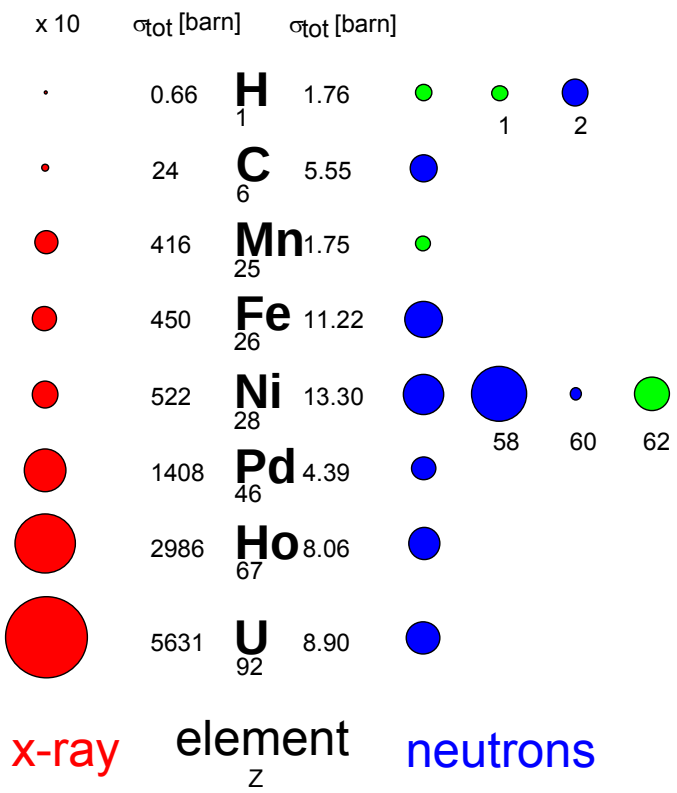


Fig. 9: *Comparison of the coherent scattering cross-sections for x-rays and neutrons for a selection of elements. The area of the circles represents the scattering cross section (x 10 for x-rays). For neutrons, the green and blue coloured circles distinguish between attractive (green, negative scattering length) and repulsive interaction (blue, positive scattering length).*

As neutrons interact strongly with the magnetic environment, polarized neutrons also show a contrast between layers with different magnetization (see section 6).

In case of low contrast, there might be a chance to enhance the contrast by contrast variation. In the case of neutrons, isotopic substitution is a good way to improve the contrast. Especially for two polymer layers, the contrast can be enhanced substantially by deuteration of one of the polymers, because the scattering lengths of Hydrogen ^{1}H and Deuterium ^{2}H differ strongly from each other (see the example in section 5.2). In the case of x-rays, one can modify the contrast for one element by tuning the x-ray energy to a resonance. An example is shown in section 5.3.

5 Layers and interfaces: Specular reflectivity of neutrons and x-rays

5.1 Neutron Reflectivity from a single Ni layer on glass

The first example is a neutron reflectivity measurement from a glass plate coated with a Ni layer. Such Ni coated glass plates are used in the neutron guides at several research reactors to guide cold neutron beams from the reactor to the instruments without losses. Fig. 10 shows the reflectivity curve together with a fit and several simulations.

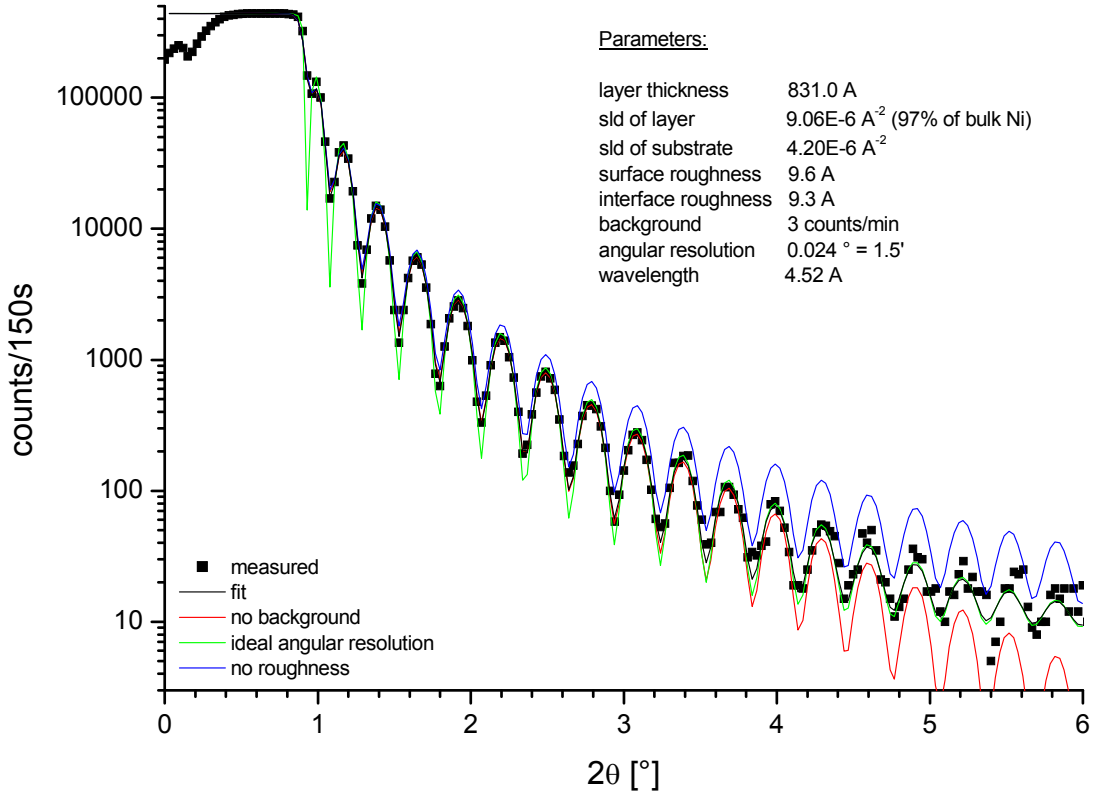


Fig. 10: *Specular reflectivity of neutrons from a Ni layer on a glass substrate, as measured with the HADAS reflectometer at the research reactor DIDO in Jülich. The black line shows the best fit, the coloured lines show simulations obtained by ignoring several parameters.*

Despite the experimental artefacts below $2\theta = 0.4°$, where the sample does not yet cover the entire beam, a typical reflectivity curve of a single layer has been measured. Up to the critical angle of $2\theta_c = 0.93°$, all impinging neutrons are reflected. At higher angles, the interference fringes of the reflections at the surface and at the interface are clearly visible. As the reflectivity drops quickly above the critical angle, it is plotted on a logarithmic scale.

It can be seen that the fit takes into account 8 parameters; some of them describe the experimental setup and some of them the sample's properties. The main parameters from the experimental setup are the wavelength used, the angular resolution of the incoming beam and the background. In the simulation of the green curve, where the resolution has not been taken into account, one can see the minima much more distinct than in the measurement. Due to the limited resolution in the real experiment, the minima are washed out due to the higher reflectivity of neutrons impinging under slightly different angles. The red curve, where the background has been omitted, shows that the signal-to-noise ratio of the instrument is strongly limited by the background, because the incoming intensity is small compared to a x-ray beamline.

The main physical parameters of the sample that have been derived from the measurement are the scattering length densities of the layer and the substrate, the roughnesses of surface and interface and, of course, the layer thickness.

The layer thickness is the easiest, it can be estimated by the distance between the maxima. The highest scattering length density (in this case of the layer) determines the critical angle θ_c, while the difference of the scattering length densities determines the height of the fringes. From the knowledge, that the layer consists of natural Ni, we have been able to determine that the density of the layer is only 97% of the bulk density, i.e. during the sputtering process a small fraction of voids has been introduced into the layer. The information about the density and the composition are not independent! Only one of the two can be derived from the reflectivity measurement.

The effect of the roughness is a drop of the reflectivity curve at high angles. The blue curve shows a simulation of the reflectivity of perfectly flat surface and interface. The growing uncertainty of the path length between two rough interfaces with growing angle of incidence leads to a decrease of interference quality with higher angles and therefore to a drop in reflectivity.

5.2 Segregation of polymer mixture under annealing

The second example shows a structure of polymer layers, where a chemical reaction at the interface between the two polymers takes place [8]. To increase the sticking between two layers of immiscible polymers, some polymer chains have a functional group ("telechelic") added which finds reaction partners in the other layer to form a multiblock copolymer (see Fig. 11). This multiblock copolymer is then fixed to the interface, with some sections compatible to polymer A and some sections compatible to polymer B.

The chemical reaction changes the sample only at the buried interface, no sign of the modification of the system is present at the surface.

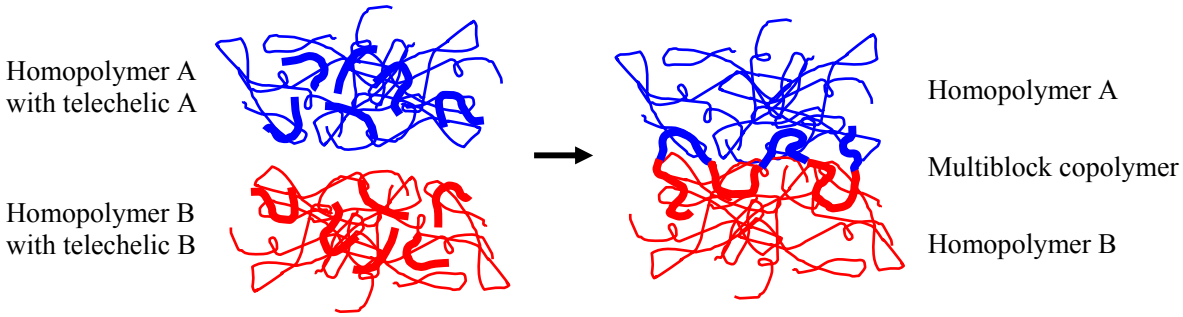


Fig. 11: *A schematic of the in situ interfacial reaction of telechelic oligomers in an immiscible polymer blend resulting in the formation of multiblock copolymer*

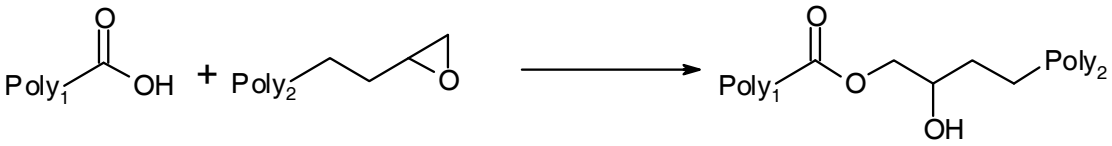
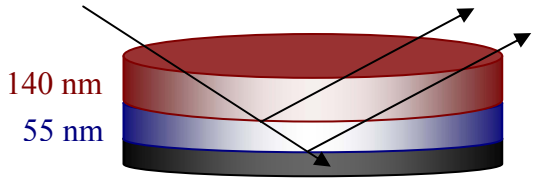


Fig. 12: *A schematic of the reaction that takes place between the carboxylic acid and epoxy groups that were present on the reactive polymer chain ends*



PS/dPS diepoxy telechelic (10%)
PMMA/PMMA-COOH (50%)
Silicon

Fig. 13: *Sample used for the neutron reflectivity experiment.*

Fig. 13 shows the sample used for this investigation. The bottom layer consists of poly methyl methacrylate (PMMA) and PMMA with a carboxylic acid group at the end. The top layer consists of polystyrene (PS) and deuterated PS with epoxy functional groups. The mobility of the polymer chains at room temperature is low, so that only the molecules react with each other that meet by chance at the interface. The reaction scheme is shown in Fig. 12. During annealing in vacuum at 150°C the mobility is increased strongly, so that more reactive groups diffuse towards the interface, where they react and become immobile.

The strong contrast between the deuterated functionalized PS and all other polymers with natural hydrogen makes the enrichment of the reactive polymer at the interface clearly observable.

Fig. 14 shows the neutron reflectivity data measured on the virgin sample as well as after several annealing steps. It can be seen that the shortest period hardly changes but the modulation of the peak height changes drastically during annealing. The analysis of the data resulted in the scattering length density profile shown in Fig. 15.

The main effect during the first 15 min annealing is a sharpening of the PS/PMMA interface and a slight increase in the density of both polymers. At the same time, the total thickness of the sample is slightly reduced. This leads to the interpretation that voids in the layers have been filled. After the following annealing steps the formation of the interface layer with a high concentration of the reactive compound can be observed due to the high scattering length density of dPS. At the same time, the remaining PS layer is depleted from dPS, so that the scattering length density of this part is reduced.

From the integrated area of dPS peak of the scattering length density profile, the amount of immobilized copolymer could be derived and compared to calculations of the diffusion velocity in the polymer melt. It was shown that the kinetics of this reaction does not match the predictions of a purely diffusion-based theory.

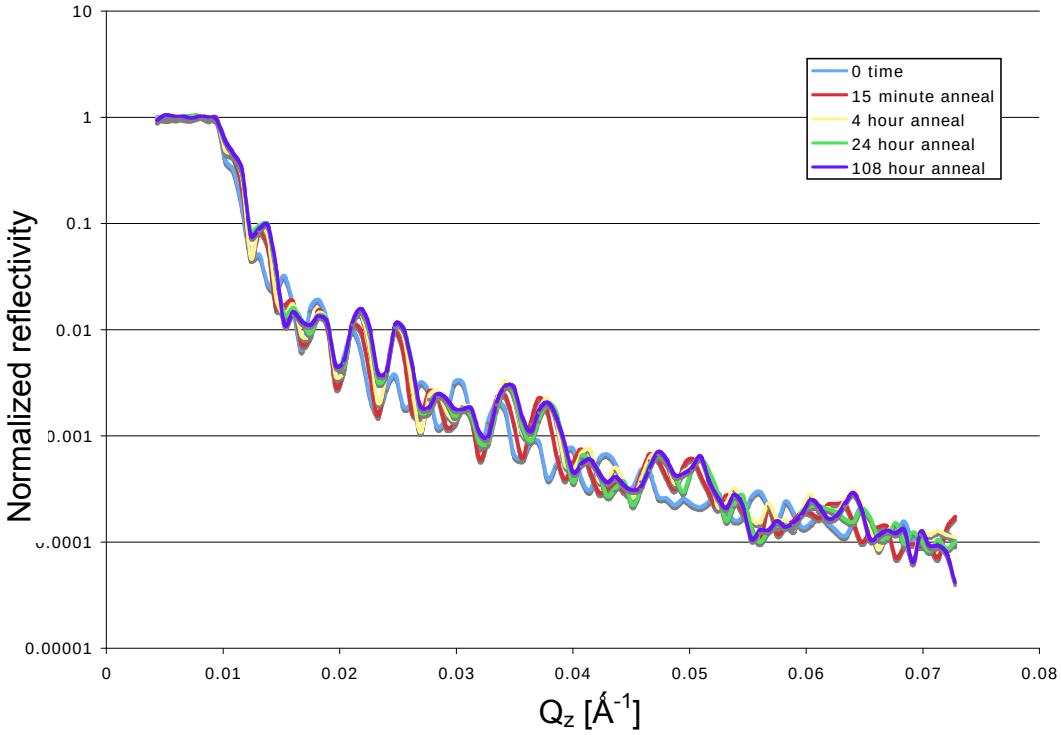


Fig. 14: *Neutron reflectivity fits from the sample shown in Fig. 13 as cast (light blue) and after 15 min, 4 hours, 24 hours and 108 hours annealing at 150 °C.*

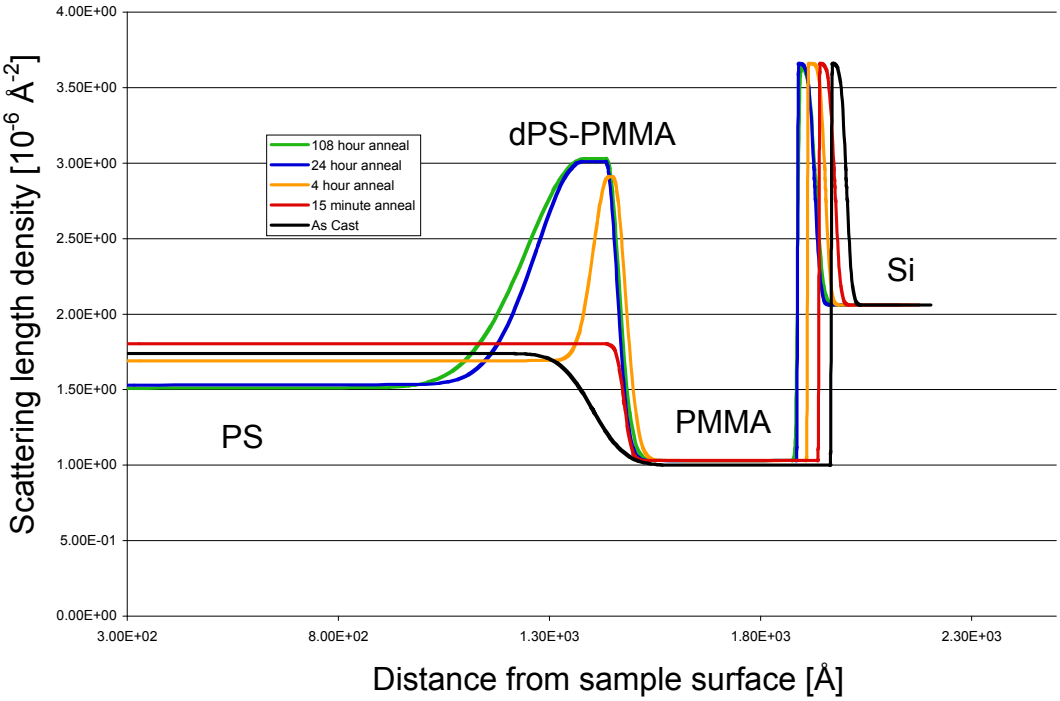


Fig. 15: *Scattering length density profile received from fitting the reflectivity data shown in Fig. 14. The sample surface is at distance 0, the Si substrate is continued towards larger distances. The colours indicate the same annealing steps as in Fig. 14*

5.3 Fe/Cr trilayer: Contrast variation with x-rays

Here, I show measurements of the buried interfaces of an epitaxially grown Fe / Cr / Fe trilayer structure on a thick Ag buffer layer. The thickness of the Ag buffer, the absorption of the x-rays over this thickness and the big roughness at the interface between Ag and the real substrate GaAs together are sufficient to be able to treat the Ag layer as substrate and to ignore everything below. The goal of this investigation is to extract information about the interface morphology of the Fe / Cr interfaces [9]. As the interface region is thin, we need to be able to measure a large Q-range, which is not possible with neutrons today. Furthermore, the influence of the magnetism in the Fe layers would have disturbed a neutron measurement.

With x-rays, we face the problem that Fe and Cr are next-nearest neighbours in the periodic system of the elements, so that the contrast for normal x-rays is very weak. The way out is anomalous x-ray scattering at a synchrotron x-ray source, where the scattering length density can be varied by choosing a x-ray energy close to the absorption edge of one of the elements.

Fig. 16 shows the dependence of the dispersion on the x-ray energy for Fe and Cr. These values have been calculated using the calculus of Cromer and Liberman [10]. It shows a strong enhancement of the contrast at the K-absorption edge of Cr, E=5989 eV and at the Fe K-edge at E=7112 eV. Unfortunately the calculation is done for free atoms and not for solids. In solids the final states are altered due to the bandstructure and the lattice, so that the dispersion calculation is not reliable over 50-100 eV above the absorption edge. Therefore, we used an energy slightly below the Cr K-edge for high contrast at the Fe / Cr interface.

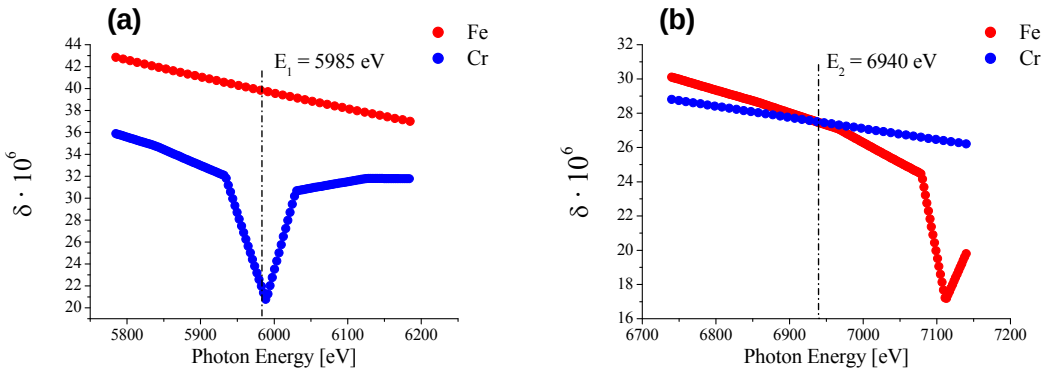


Fig. 16: *Calculated X-ray dispersion close to the Cr K-absorption edge (a) and close to the Fe K-edge (b)*

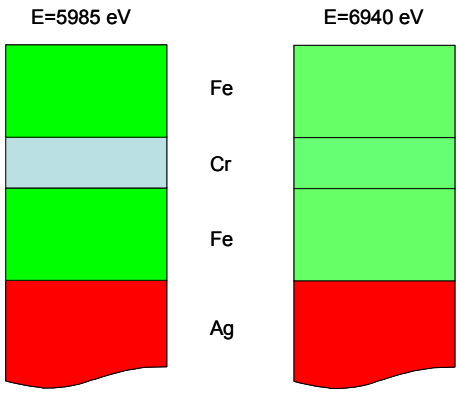


Fig. 17: *Colour schematic of the contrast in the Fe / Cr / Fe / Ag system at highest and lowest contrast*

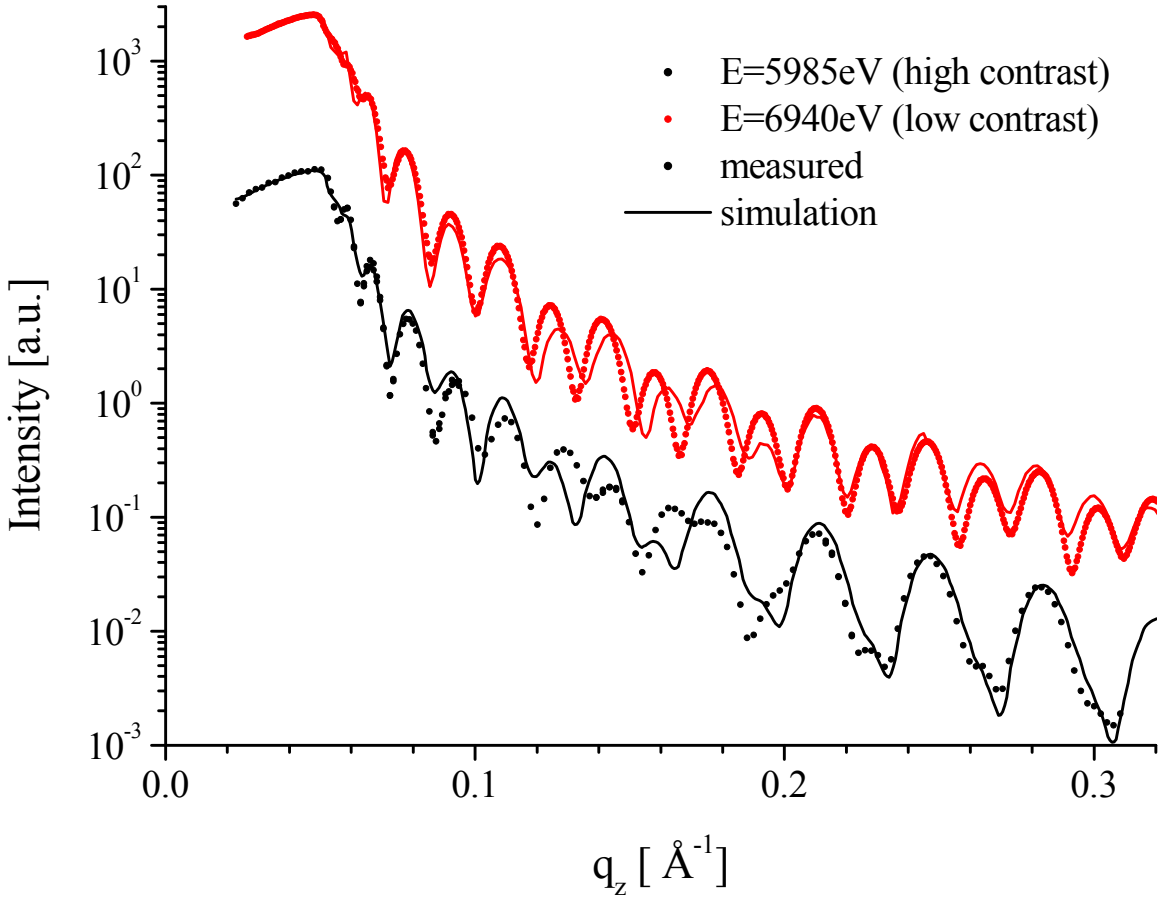


Fig. 18: *Specular reflectivity of the Fe / Cr / Fe / Ag layered structure at the two different energies with high and vanishing contrast between Fe and Cr*

At E=6940 eV, the dispersion curves for Fe and Cr cross when approaching the Fe K-edge. The contrast is vanishing. Fig. 17 shows a colour schematic of the contrasts at the two energies chosen. At 6940 eV there is no contrast at the Fe / Cr interfaces, the x-rays are reflected only at the surface and the Fe / Ag interface. The data is shown in Fig. 18. The period of the reflectivity fringes is associated with the total thickness of the Fe / Cr / Fe stack. The imperfection of the fit at about $q_z = 0.13$ Å^{-1} is due to the oxide layer on top, which is not perfectly described by the model used. The measurement at this energy can be used to fix the parameters of those interfaces which are less interesting.

At E=5985 eV, there is a strong contrast between Fe and Cr, but it is still weaker than the contrast at the surface and the contrast against the Ag substrate. The reflectivity curve is now dominated by fringes with the period according to the bilayer thickness Fe / Cr. This measurement can give reliable information about the roughness of the interface between Fe and Cr.

6 Magnetization depth profiles: Specular reflectivity of polarized neutrons

In this chapter, I would like to have a closer look at the investigation of the magnetization depth profile in magnetic layered structures. Polarized neutrons are a suitable probe for the layers' magnetization, because the neutron is a spin ½ particle and therefore interacts with the magnetic induction **B**. As reflectometry measurements average over the atomic structure of the sample, antiferromagnetic structures are not accessible due to the vanishing net magnetization. In contrast to that, magnetization densities of ferromagnetic layers can be measured on an absolute scale with reflectometry of polarized neutrons.

6.1 Index of refraction of magnetized material for polarized neutrons

We stay in the approximation of homogeneous layers with flat interfaces, where the potential for the neutron only depends on the z coordinate. In the case of magnetic multilayers, we need to take into account the interaction of the neutron's spin with the magnetic induction inside matter. To treat this properly, we have to work with wave functions in the 2-dimensional quantum mechanical spin space, where the usual space-dependent functions, e.g. the potential, become operators on the neutron's spin.

The potential for the interaction of the neutron with matter in the layer l can be separated into two parts

$$\hat{V}_l = V_l^N \hat{1} + \hat{V}_l^M \tag{25}$$

where V_l^N is the nuclear interaction from eq. (3), $\hat{1}$ is the unity operator, which does not affect the spin state, and $\hat{V}_l^M = -\mu_n \hat{\boldsymbol{\sigma}} \cdot \mathbf{B}_l$ is the magnetic dipole interaction operator between the neutron magnetic moment operator $\mu_n \hat{\boldsymbol{\sigma}}$ and the magnetic induction $\mathbf{B}_l$.

We assume the direction of the external magnetic field $\mathbf{H}=H\mathbf{e}_x$ to be oriented in the x-direction of the coordinate system defined in Fig. 5. Then it is convenient to choose also the x-axis as quantization axis for the neutron spin, so that the order of the Pauli matrices in the spin operator $\hat{\boldsymbol{\sigma}} = (\sigma_x, \sigma_y, \sigma_z)$ is the following.

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{26}$$

It is arbitrary to choose the coordinate system for the scattering experiment (Fig. 5) and for the neutron spin to be parallel; the result does not change, if the magnetic field **H**, which defines the quantization axis for the spin, is rotated to any other direction in the x-y-plane of the scattering experiment. This rotation will only change the phase of the spin-flip scattering (eq. 30), which has no influence on the reflectivities and transmissivities.

The magnetic induction can be decomposed into terms of applied field and magnetization

$$\mathbf{B} = \mu_0(\mathrm{H}\mathbf{e}_x + (1-\mathrm{D})\mathbf{M}) = \mathbf{B}_0 + \mu_0(1-\mathrm{D})\mathbf{M} \tag{27}$$

where $\mathbf{B}_0$ is the induction from the external magnetic field, D is the demagnetizing factor and $\mathbf{M}$ the magnetization of the material. In the case of a thin film, (1 – D)$\mathbf{M}$ is equal to the in-plane component of the magnetization.

The magnetization component M_z perpendicular to the film surface cannot induce any magnetic contrast between adjacent layers, because $\nabla \cdot \mathbf{B} = 0$ does not allow B_z to change discontinuously when crossing an interface. In fact, this is the origin of the demagnetizing effect.

$\mathbf{B}_0$ is constant over the sample volume and therefore gives a constant contribution to the index of refraction as well for vacuum as for every material involved. Therefore, all these contributions cancel out when calculating reflectivity R and transmissivity T according to (18) and (19).

The only remaining contributions are M_x parallel to the quantization axis and M_y perpendicular to the quantization axis, but in plane. With that, we can rewrite the total interaction operator in analogy to (5) as

$$\hat{\mathrm{V}}_l = \frac{2\pi\hbar^2}{\mathrm{m}}(\rho_l^{\mathrm{N}}\hat{1} + \rho_l^{\mathrm{M}}\hat{\boldsymbol{\sigma}} \cdot \mathbf{b}_l) \tag{28}$$

with the nuclear scattering length density ρ_l^{N}, the magnetic scattering length density ρ_l^{M} and the unit vector $\mathbf{b}_l$ along the magnetic induction vector $\mathbf{B}_l$ in layer l. In most cases, if $\mu_0 M >> B_0$, $\mathbf{b}_l$ is approximately parallel to $\mathbf{m}_l$.

In complete analogy to section 4, the Schrödinger equation can be solved in coordinate and spin space. The eigenvectors $|+\rangle$ and $|-\rangle$ of the operator $\hat{\boldsymbol{\sigma}} \cdot \mathbf{b}_0 = \sigma_x$ with the eigenvalues +1 and -1, respectively, define states of the neutron with “spin up” and “spin down”. The solution of the Schrödinger equation is the neutron state $|\Psi(\mathbf{r})\rangle$, which is again a linear combination of those two eigenvectors:

$$|\Psi(\mathbf{r})\rangle = \Psi^+(\mathbf{r})|+\rangle + \Psi^-(\mathbf{r})|-\rangle = \begin{pmatrix} \Psi^+(\mathbf{r}) \\ \Psi^-(\mathbf{r}) \end{pmatrix} \tag{29}$$

We end up with a set of two coupled one-dimensional linear differential equations for every layer:

$$\begin{aligned}\Psi_l^{+\,\prime\prime}(z)+\left[k_{zl}^2-4\pi(\rho_l^N+\rho_l^M m_{xl})\right]\Psi_l^+(z)-4\pi\rho_l^M m_{yl}\Psi_l^-(z)=0\\ \Psi_l^{-\,\prime\prime}(z)+\left[k_{zl}^2-4\pi(\rho_l^N-\rho_l^M m_{xl})\right]\Psi_l^-(z)-4\pi\rho_l^M m_{yl}\Psi_l^+(z)=0\end{aligned} \tag{30}$$

The solution of this set of differential equations can be done in analogy to (23). The calculation is straightforward, but lengthy. Its solution can be found in [11]. It ends with the reflection and transmission operators $\hat{r}_l$ and $\hat{t}_l$, which are again operators in spin space with two different eigenvalues.

Four types of reflectivities can be measured, which are the squares of the projections of the reflection amplitude onto spin up or down neutron states:

$$R^{\pm\pm}=\left|\langle\pm|\hat{r}_0|\pm\rangle\right|^2 \tag{31}$$

These reflectivities have a clear physical meaning. From eq. (30) it is clear that only nuclear scattering and the magnetization component M_x parallel to the field lead to non-spinflip (NSF) reflectivities. By analysis of the sum and the difference of R^{++} and R^{--}, these two contributions can be separated from each other. Fig. 19 shows the influence of the magnetic scattering length density on the critical angle and the Reflectivity for the case of a magnetically saturated surface, where only NSF reflectivity is present.

The spin-flip (SF) reflectivities R^{+-} and R^{-+} are equal to each other and arise from the magnetization component M_y in plane, perpendicular to the field. As the sign of the magnetization is only coded in the phase of the SF reflected wave, it is lost when measuring the intensities. Therefore, only $|M_y|$ can be measured.

In conclusion, polarized neutron reflectivity can measure the nuclear scattering length density ρ^N and the magnetization components M_x and $|M_y|$ for every layer in the thin film structure.

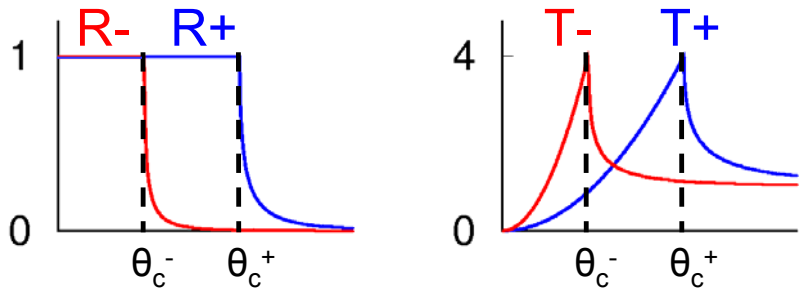


Fig. 19: *Reflectivity and Transmissivity for polarized neutrons from a saturated ferromagnetic material*

6.2 Magnetic contrast: Co/Cu multilayer

This example shows experimentally the contrast for polarized neutrons which depends as well on the nuclear as on the magnetic scattering length densities. We have measured the reflectivity of a multilayer with 20 periods of a Co layer and a Cu layer [12]. The thickness of the Cu layer is so that the Co layers are coupled ferromagnetically, and the sample has been saturated in a strong magnetic field. During the experiment, the guiding field, which defines the quantization axis for the neutron spin, is kept parallel to the magnetization of the Co layers.

The nuclear scattering length density for Co is 2.3 E-6 $Å^{-2}$. For Cu it is much higher 6.53 E-6 $Å^{-2}$. The magnetic scattering length density for Co is 4.24 E-6 $Å^{-2}$. So the sum of nuclear and magnetic scattering length density of Co hardly differs from the scattering length density of the non-magnetic Cu, so that there is no contrast for spin-up neutrons, i.e. neutrons with the spin parallel to the external field. The contrast for spin-down neutrons, which feel the difference of nuclear and magnetic scattering length density, is huge. Fig. 20 shows a colour sketch of the different contrasts.

Fig. 21 shows the reflectivity curve for spin-up and spin-down neutrons. As no magnetization component perpendicular to the field is present, the spin-flip channels only contain a signal coming from the imperfect polarisation of the beam and are not shown. The total reflection angle for both channels is the same because it is dominated by Cu. Due to the large contrast for spin-down neutrons, two Bragg peaks at $2\theta=3°$ and at $2\theta=6°$ corresponding to the structural periodicity are easily observed already after short beam time. The reflectivity curve for spin-up neutrons is rather unstructured due to the missing contrast between magnetized Co and Cu.

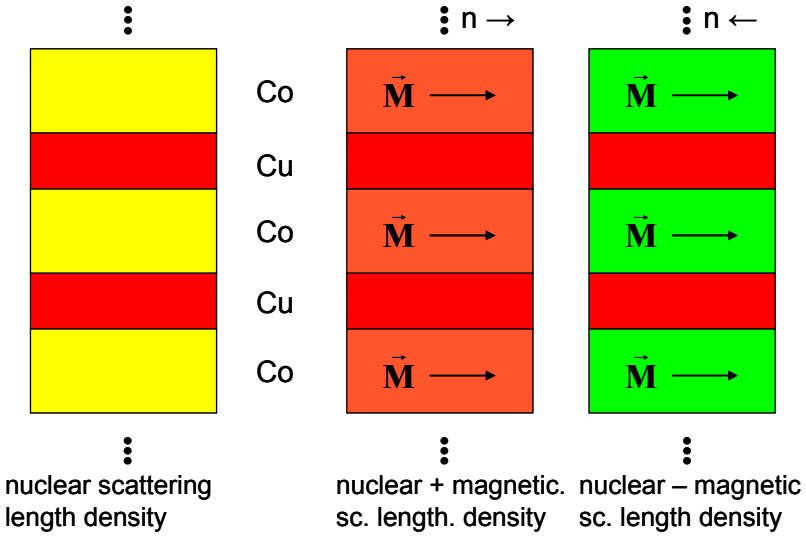


Fig. 20: *Colour representation of the contrasts in a magnetically saturated Co / Cu multilayer. Purely nuclear contrast (left), contrast for spin up neutrons (middle) and contrast for spin down neutrons (right)*

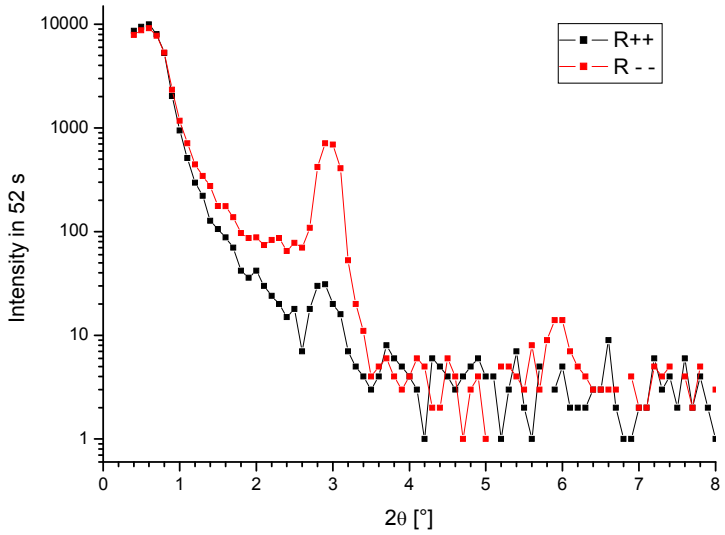


Fig. 21: *Specular reflectivity of polarized neutrons from a magnetically saturated Co/Cu multilayer. For spin-down neutrons two Bragg peaks at 2θ=3° and 6° are clearly visible while the reflectivity curve for spin-up neutrons is hardly structured.*

6.3 Layer-by-layer magnetometry: Polarizing supermirror

Polarizing supermirrors are commonly used for the polarization of cold neutron beams. For the operation of a polarized neutron instrument, it is of course important to know the performance of the polarizers used and their behaviour in the magnetic field. During the characterization of the device, it turned out that a lot of physics of the layered magnetic structure can be understood by having a closer look at the data, and so we have investigated the polarizing supermirror thoroughly [13,14]. In this section, I will present the laterally averaged data from specular reflectivity; the off-specular and GISANS measurements can be found in reference [14].

A polarizing supermirror is a stack of bilayers of a magnetic and a non-magnetic material. The thickness of the bilayers is gradually increasing to receive constructive interference for a broad range of Q-values, what results in an extended plateau of high reflectivity above the total reflection angle. The materials are chosen to have a high contrast for one spin direction and no contast for the other.

The supermirror we have investigated is a polarizing supermirror produced by Swissneutronics which reflects the neutrons with the proper spin direction up to m=2.5 times the critical angle of Ni. The first three columns of Fig. 22 show the intended scattering length density sequence. 100 bilayers of FeCoV (ferromagnetic) and TiN (nonmagnetic) with gradually increasing thickness are transparent for neutrons with spin antiparallel to the

magnetization and reflecting for neutrons with spin parallel to the magnetization. Below the stack, there is a strongly absorbing Gd layer to absorb all spin down neutrons and a glass substrate for mechanical stability.

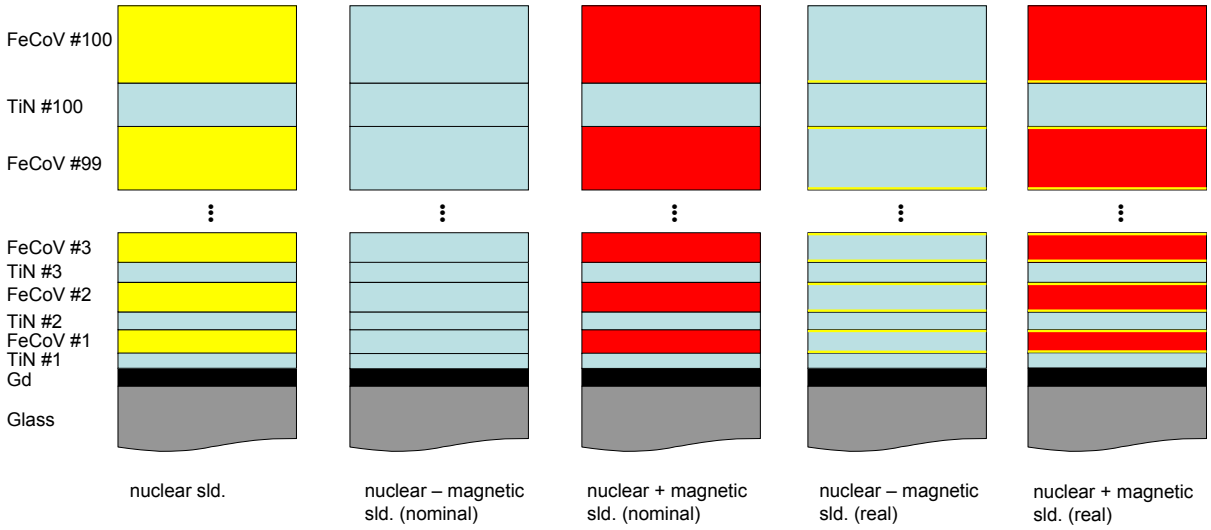


Fig. 22: *Colour representation of the scattering length density distribution in a polarizing supermirror. On the left, the nuclear scattering length density distribution is shown. The next two columns show the intended contrast for neutrons with spin antiparallel to the magnetization (no contrast) and for neutrons with spin parallel to the magnetization (high contrast). The two columns on the right show the measured scattering length density profile. We found magnetically dead layers at the interfaces between FeCoV and TiN, so that a 2-4Å of the FeCoV layer show only the nuclear scattering length density, but no magnetic contribution.*

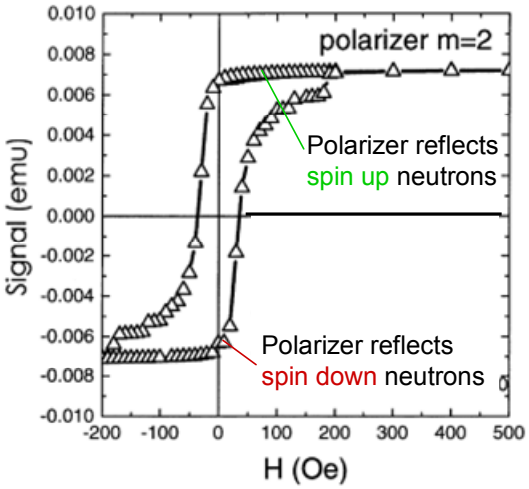


Fig. 23: *Magnetization curve of a polarizing supermirror with 40 bilayers [15]*

As our sample cannot be measured in a magnetometer because the back surface is also coated magnetically, we show the magnetization curve of a comparable sample with only 40 bilayers in Fig. 23. After negative saturation, the polarizer remains almost completely saturated in negative direction at small positive fields. This allows two different working conditions for the polarizer. On the one hand, it can work in saturation, i.e. after exposure to a high positive magnetic field. Then all magnetic layers are magnetized along the field direction, and the supermirror reflects spin-up neutrons, i.e. neutrons with spin along the magnetic field. On the other hand, it can be used in remanence with all layers magnetized opposite to the applied (small) magnetic field. Then the spin-down neutrons are reflected.

Fig. 24 shows the polarized neutron reflectivity measurement at different fields after negative saturation. Fig. 24 a) shows the remanence. In the "working range" up to α_i = 21 mrad, the reflectivity for spin down neutrons R^{--} is more than two orders of magnitude higher than R^{++}, yielding a good negative polarization of the reflected beam. The reflectivity for the proper polarization channel is close to 1, the increasing slope at small angles is an experimental artefact due to the increasing coverage of the beam's cross section with increasing angle of incidence. The spin-flip signal is completely determined by the imperfect polarization of the incident beam, no spin-flip specular reflectivity is coming from the sample. Fig. 24 d) is measured at saturation and is almost exactly opposite to the remanent state.

In the two intermediate states shown in Fig. 24 b) and c), there is a crossover between R^{++} and R^{--}, showing that some of the layers are magnetized along and some antiparallel to the field. In the fit, we can address the individual magnetization of every layer and find out, which layers have flipped and which stay magnetized opposite to the field direction. It turned out, that the thinner, i.e. lower magnetic layers flip first. In the case of 3.8 mT (Fig. 24 b), 48 layers have flipped, in the case of 5.6 mT (Fig. 24 c), 94 FeCoV layers have flipped in field direction. Fig. 25 shows the number of layers flipped in field direction as a function of the field strength.
What is striking, is the relatively high R^{--} signal in saturation that drops at twice the angle of the end of the supermirror plateau. It shows that spin-down neutrons see a structure that has half the period of the intended supermirror structure for the spin-up neutrons. It turned out, that this feature comes from magnetically dead layers at the interface between FeCoV and TiN. At the top and the bottom of the FeCoV layer, there is a 2 – 4 Å thick nonmagnetic region with the purely nuclear scattering length density of FeCoV. The two pictures on the right side of Fig. 22 show a schematic of the real scattering length density profile, where the reflectivity in R^{--} comes from the yellow – blue bilayers, that have about half the thickness of the red – blue bilayers that result in R^{++}.

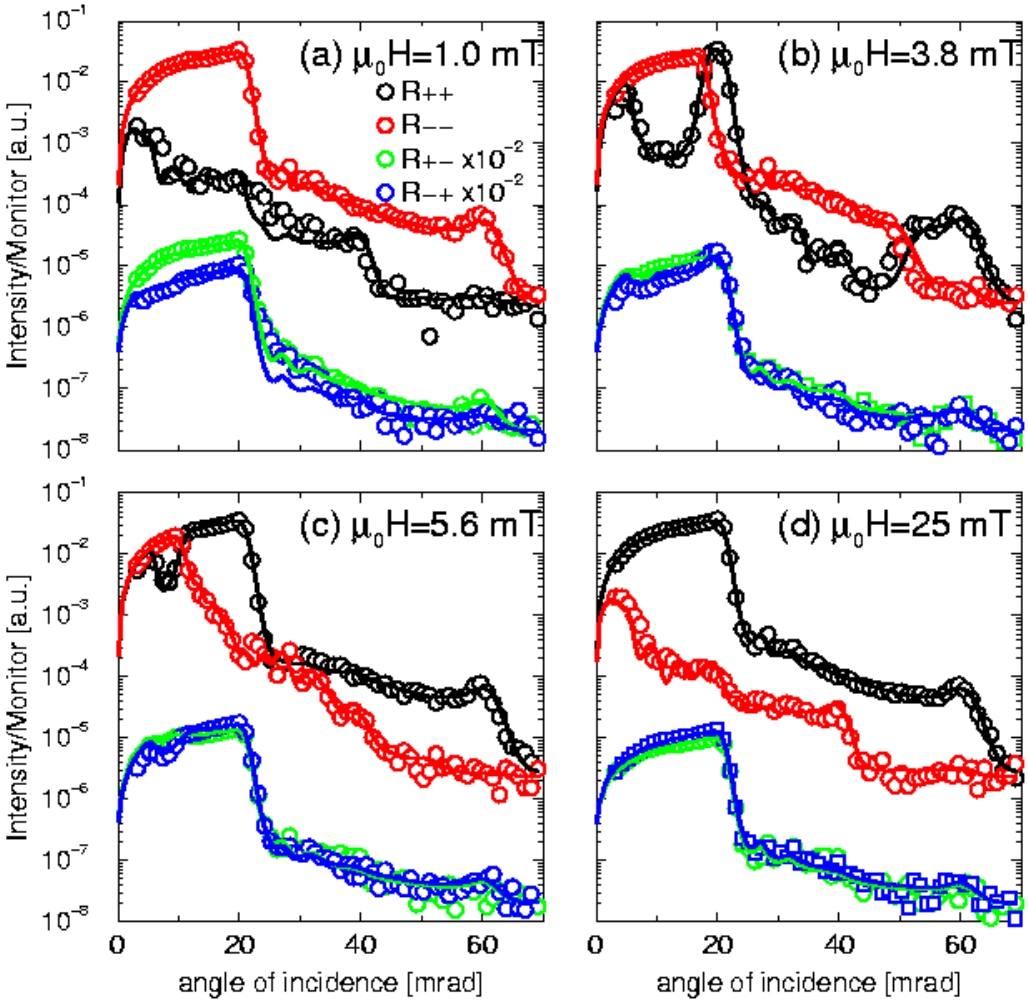


Fig. 24: *Polarized neutron reflectivity measurements on the polarizing supermirror at different magnetic fields applied after negative saturation. The dots show experimental data, the lines are fits.*

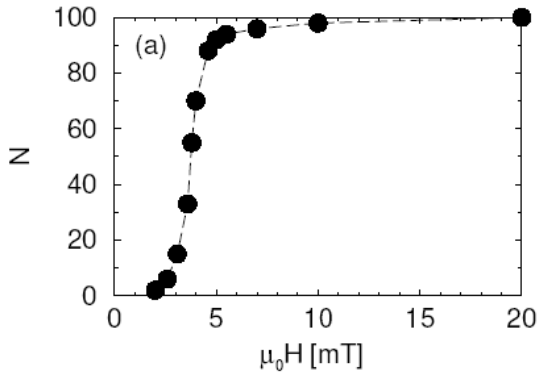


Fig. 25: *Number of FeCoV layers flipped in field direction as a function of the positive magnetic field applied after negative saturation*

7 Scattering from lateral fluctuations

7.1 Lateral correlations in layered structures

The normal case of real layered structures is that they are not completely laterally homogeneous but they show at least some roughness at the interfaces. Chemical deviations may come from interdiffusion at an interface or oxidation processes at the sample surface. Magnetic layers split laterally into magnetic domains which are visible for polarized neutrons. Lateral structuring of layered structures is of high technological importance for many applications in modern electronic devices.

If lateral fluctuations (i.e. all local variations in scattering length density from the laterally averaged scattering length density) exist with a correlation length smaller than the projected coherence lengths (eq. (2) and (3)), then scattering under grazing incidence can be observed. At grazing incidence, those coherence lengths can extend into the µm range and the interaction of the x-ray photons or neutrons with the fluctuations will lead to multiple scattering processes within this surface, so that the Born approximation will not be able to describe the effects properly.

In most cases, the layered structure is dominant, so that the main scattering feature is the specular reflection that is caused by the laterally averaged scattering length density contrasts. In this case, the potential inside layer l can be decomposed into two terms describing the laterally averaged potential (as main potential $V_l^0(z)$) and the fluctuations around the main potential $V_l^1(\mathbf{r})$.

$$V_l(\mathbf{r}) = V_l^0(z) + V_l^1(\mathbf{r}) \tag{32}$$

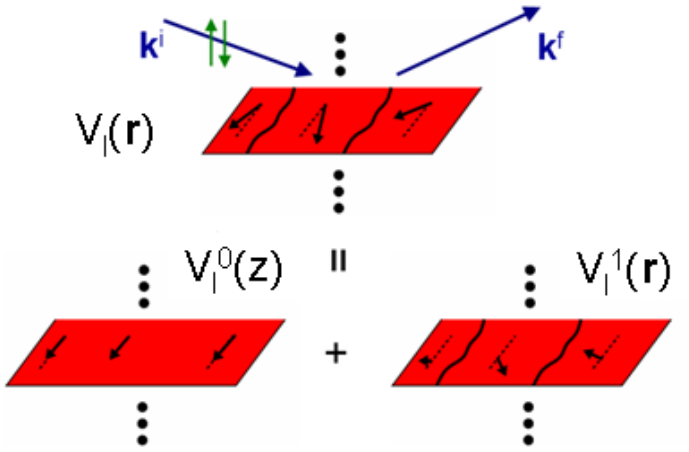


Fig. 26: *Decomposition of the potential of layer l for the treatment with the DWBA*

Fig. 26 shows the decomposition of the potential of the layer no. l in the case of magnetic domains. The main potential V^0 shows a homogeneous magnetization for every layer. It is defined by the lateral average magnetization direction and modulus. V^1 contains all deviations in magnetization direction and the domain walls.

In a comparable way, the formalism can be defined for interface roughness, interdiffusion or chemical precipitations.

The main potential then can be treated exactly (see section 4), and the fluctuations can be treated in the framework of the Distorted Wave Born Approximation (DWBA) as a perturbation. Then, all refraction effects close to θ_c and reflected wavefields are taken into account as origin of off-specular scattering. Depending on the length scale of the correlations, they can be seen in offspecular scattering (in x-direction, where the coherence length extends into the μm range) or in GISAS geometry for fluctuations in the nm scale.

A description of the DWBA formalism for the case of magnetic domains can be found in [16]. Different cross sections are used in the case for interface fluctuations or density fluctuations within layers separated by flat interfaces [17, 18, 19, 20, 21, 22]. Today, off-specular scattering from many structures can be modelled successfully. Modelling GISAS scattering is still a challenge in the framework of the DWBA, but first successful results are published for systems with a high degree of order.

7.2 Off-specular scattering from nanoparticle supercrystals

Here, I would like to present the scattering results from a system, where very different length scales are realized in a single sample. The sample consists of spherical Fe_2O_3 nanoparticles with 9.5 nm diameter surrounded by an organic ligand shell, which originally are dispersed in an organic solvent. When this solution is drop-casted onto a flat Si wafer surface and the solvent evaporates slowly, the nanoparticles self-assemble in well-ordered supercrystals with dimensions up to 2 μm in diameter and 1 μm in height. These supercrystals are then randomly spread over the wafer surface [23, 24] with a typical spacing of a few μm.

Fig. 27 shows a scanning electron microscope (SEM) picture of the arrangement of supercrystals on the surface of a Si wafer. One can see a lot of crystals which are bigger than 1 μm and spread over the surface with several μm of space in between. In between the supercrystals, some single nanoparticles remain on the Si surface, which are not ordered. When zooming in, one can see the perfect order of the nanoparticles within a single supercrystal (Fig. 28). The SEM picture gives the impression that the nanoparticle layers inside the supercrystals are parallel to the surface of the Si wafer.

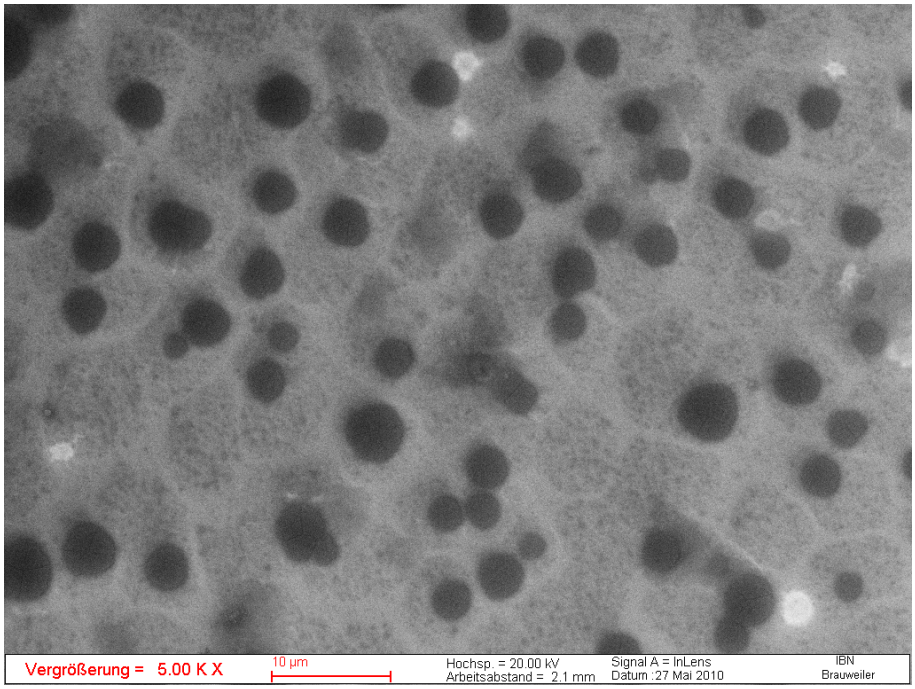


Fig. 27: *SEM (Scanning electron microscope) picture of self-organized nanoparticle supercrystals dispersed over the surface of a Si wafer.*

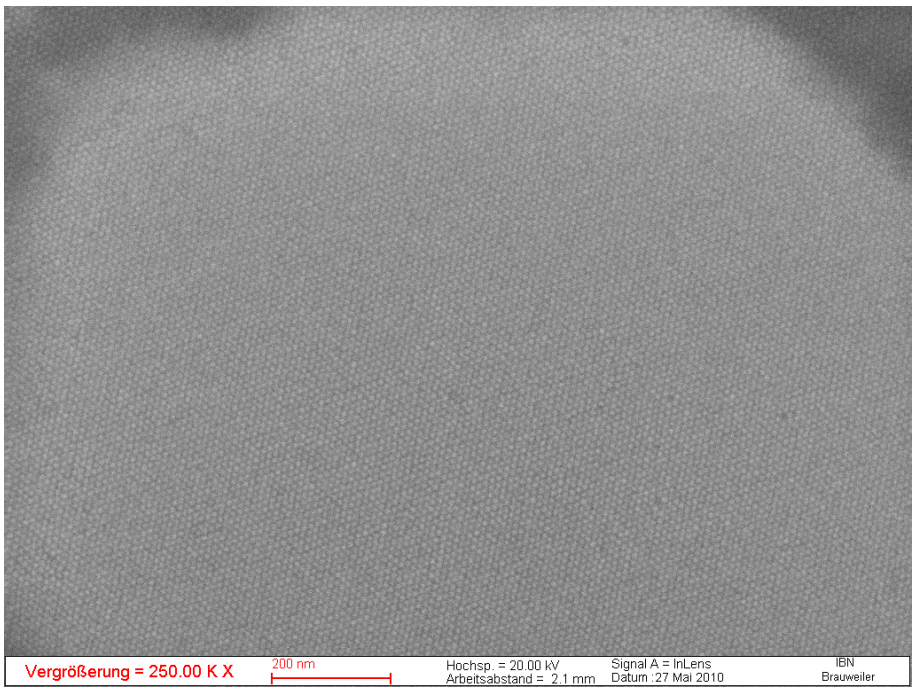


Fig. 28: *SEM picture zoomed into one supercrystal showing the perfect order of the nanoparticles on the surface of the supercrystal*

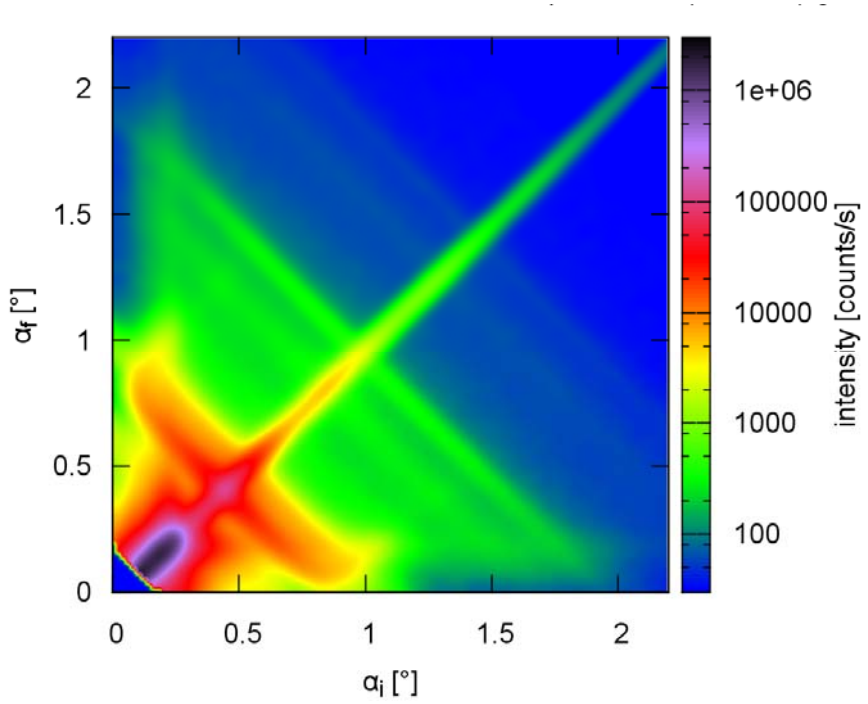


Fig. 29: *Specular reflectivity and offspecular x-ray scattering from an array of nanoparticle supercrystals. The specular reflectivity can be found in the main diagonal of the image, where* $\alpha_i = \alpha_f$.

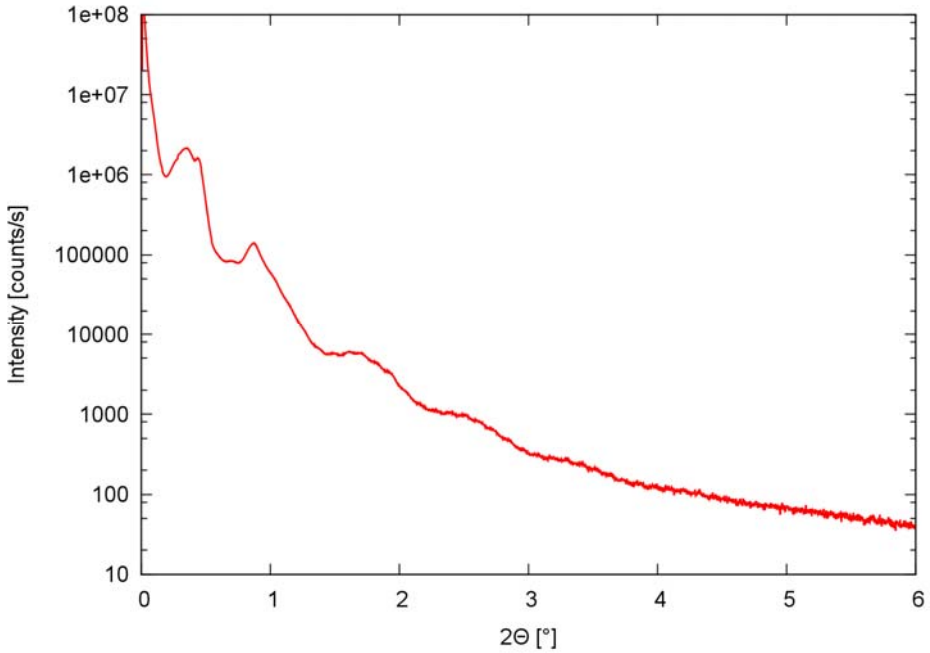


Fig. 30: *Specular reflectivity extracted form the measurement presented in Fig. 29*

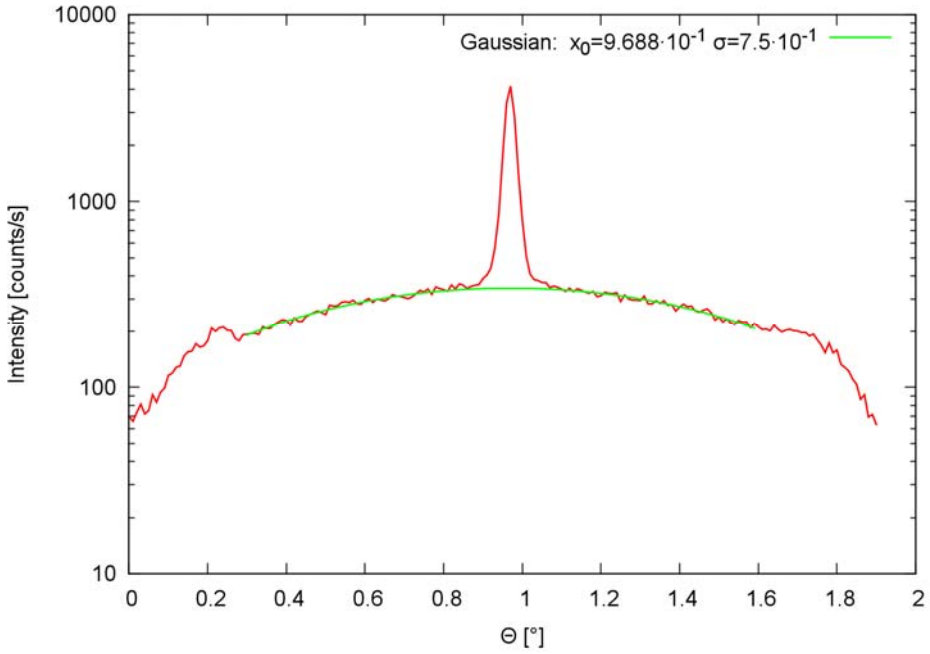


Fig. 31: *Rocking scan at* $\alpha_i + \alpha_f = 1.9°$ *out of the measurement presented in Fig. 29. In the middle, one can see the specular reflection which dominates the curve. Below the specular reflection there is offspecular scattering which can be fitted with a Gaussian describing the lateral correlations between the neighboring supercrystals. The peaks at* $\theta = 0.25°$ *and* $\theta = 1.65°$ *are due to the Yoneda effect at* $\alpha_i = \theta_c$ or $\alpha_f = \theta_c$, resp.

Fig. 29 shows the specular reflectivity and the off-specular scattering measured from the surface of such a supercrystal array with x-rays with Cu Kα wavelength (λ = 1.54 Å). One can see the specular reflectivity dominant in the main diagonal of the figure together with a lot of offspecular scattering, mainly perpendicular to the main diagonal.

The specular reflectivity, which is extracted in Fig. 30, shows clear oscillations showing a layered structure perpendicular to the sample surface. This is the first proof of the alignment of the inner structure of all (!) supercrystals with the surface of the Si wafer. The period of the oscillations reveals the layer thickness, which in this case is the thickness of a close-packed layer of nanoparticles with organic shell. Therefore, the layer thickness is slightly thicker than the nanoparticle diameter.

In offspecular scattering, one can observe many Bragg sheets perpendicular to the specular direction. These Bragg sheets arise from lateral fluctuations which are periodical in the thickness. In this case, they are the supercrystals which obviously have the same place in every layer. Close to the axes of the plot in Fig. 29 one can see that the offspecular scattering features are bent due to refraction effects. The intensity is enhanced there because of the increased wavefield along the surface close to the critical angle θ_c (Yoneda effect, cf. Fig. 8).

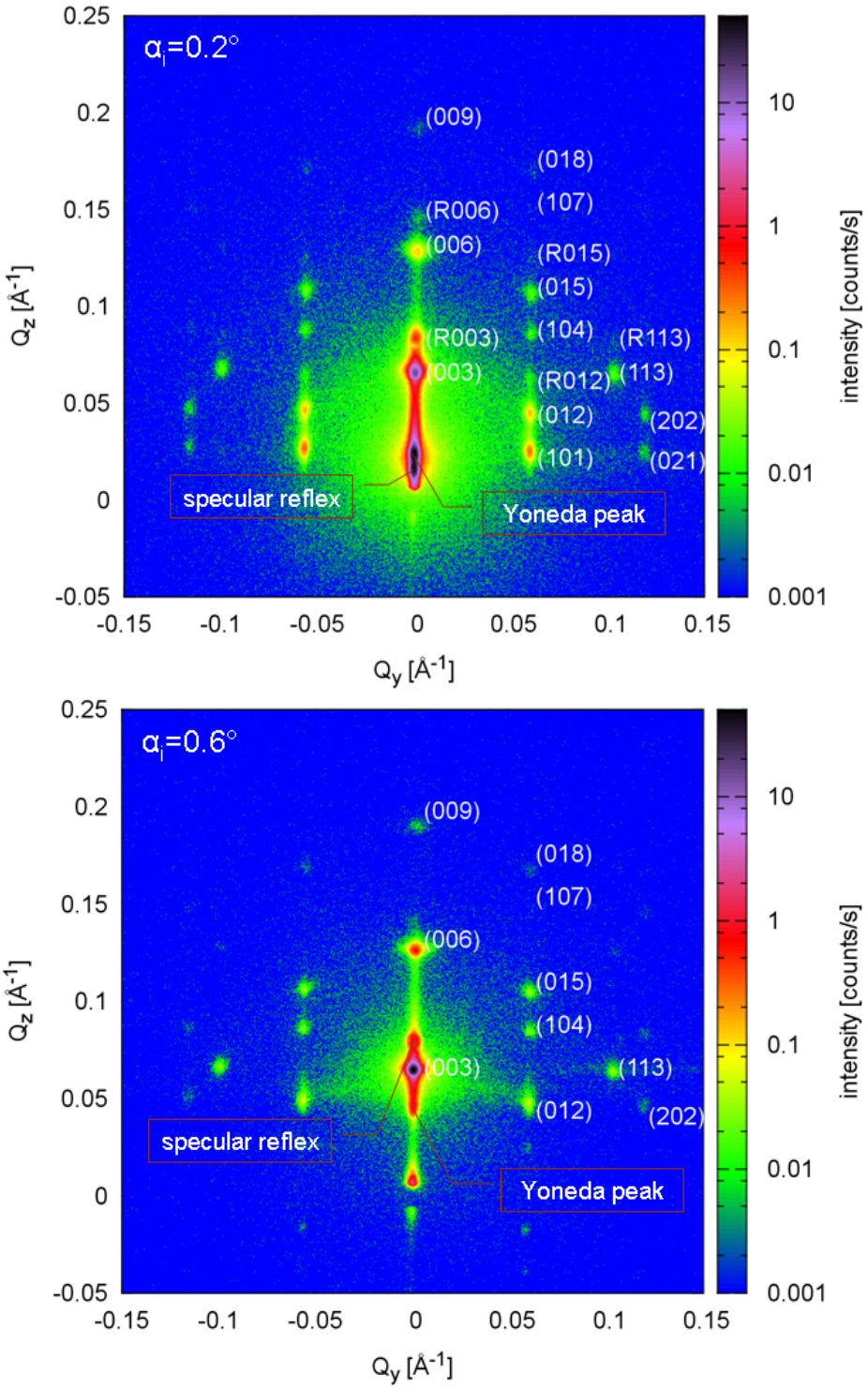


Fig. 32: *GISAXS scattering from the array of nanoparticle supercrystals at two different angles of incidence: 0.2° (top) slightly below the total reflection angle* θ_c *and 0.6° (bottom) well above* θ_c*. The diffraction spots are indexed, the peaks with index R come from diffraction of the reflected beam.*

Fig. 31 shows a cut perpendicular to the specular direction at $\alpha_i + \alpha_f = 1.9°$. In the centre, the specular reflectivity clearly dominates the curve. At the critical angles $\alpha_i = \theta_c$ and $\alpha_f = \theta_c$ the offspecular scattering is increased due to the Yoneda effect. In the range between, one can nicely fit the offspecular scattering with a Gaussian curve. The width of the Gaussian here gives the average periodicity of the lateral fluctuations, in this case the distance between the centres of the supercrystals. In this case, it turns out to be an average periodicity of 13.6 µm. When comparing to Fig. 27, one has to take into account that the coherence area is long in x-direction but narrow in y-direction, so that this periodicity only resembles the correlations along the x-direction. Neighbours that may be closer in y-direction are "invisible".

Fig. 32 shows the GISAXS pattern at two different angles of incidence. One can see a lot of diffraction peaks arising from the order of the nanoparticles inside the supercrystals. As there are thousands of supercrystals on the wafer, all in-plane orientations are present, so that the diffraction pattern observed is an average of a 2-dimensional powder-like arrangement. Only in the direction perpendicular to the surface, the supercrystals are aligned to each other. Due to this fact, e.g. the reflections (104) and (015) can be observed simultaneously, although they need a different orientation of the supercrystal in the x-ray beam. Indexing the diffraction spots yields the trigonal crystal structure $R\bar{3}$ with a lattice parameter again close to the nanoparticles' diameter.

Together with the diffraction spots one can observe the specular reflection and the Yoneda peak, both at $Q_y = 0$. At the Yoneda angle $\alpha_f = \theta_c$ there is as well increased scattering from periodicities parallel to the surface due to the enhanced wavefield, as it is observed in off-specular scattering. Especially in the bottom graph in Fig. 32 one can see the line of enhanced scattering horizontally through the Yoneda peak.

If the angle of incidence is sufficiently small, so that the reflectivity is close to 1, the diffraction from the reflected beam is comparably strong as the diffraction from the incoming beam. In this case, additional diffraction spots are visible which move with changing α_i, while those arising from the incoming beam don't move. In the top graph in Fig. 32 these peaks are indexed with "R".

It is obvious, that here we observe the inner structure of the supercrystals in the range of 10 nm with GISAXS in the y-direction perpendicular to the scattering plane, while before we looked at the arrangement of the supercrystals in the range of 10 µm with off-specular scattering in the x-direction in the scattering plane. According to eq. (1) these different ranges of lateral periodicities are accessible due to the anisotropy of the arrangement between sample surface and beam direction.

References

[1] P. Grünberg, J. Phys. Cond. Matter **13**, 7691 (2001)
[2] Physik Journal **5**, Nr. 11, p. 17 (Nov. 2006)
[3] U. Nowak, K.D. Usadel, J. Keller, P. Miltényi, B. Beschoten, G. Güntherodt, Phys. Rev. B **66**, 014430 (2002)
[4] H. Oldag, A. Scholl, F. Nolting, E. Arenholz, S. Maat, A.T. Young, M. Carey, J. Stohr, Phys. Rev. Lett. **91**, 017203 (2003)
[5] I.V. Roschin, O. Petracic, R. Morales, Z.-P. Li, X. Battle, I.K. Schuller, Europhys. Lett. **71**, 297 (2005)
[6] H. Frielinghaus, Lecture D1 of this book
[7] S. Blügel, Lecture A2 of this book
[8] J.K. Rice, "The reaction of telechelic polymers at multicomponent interfaces: A molecular loop study", Dissertation, University of Tennessee, Knoxville, USA, Dec. 2006
[9] M. Feygenson, E. Kentzinger, N. Ziegenhagen, U. Rücker, G. Goerigk, Y.G. Wang and Th. Brückel, J. Appl. Cryst. **40**, 532 (2007)
[10] D.T. Cromer, D. Liberman, J. Chem. Phys. **53**, 1891 (1970).
[11] E. Kentzinger, U. Rücker, B. Toperverg, Physica B **335**, 82 (2003)
[12] W. Babik, U. Rücker, Experimental Report FRJ-2 2002, p. 167
[13] U. Rücker, E. Kentzinger, B. Toperverg, F. Ott, T. Brückel, Appl. Phys. A **74**, S607
[14] E. Kentzinger, U. Rücker, B. Toperverg, F. Ott, Th. Brückel, F. Ott, Phys. Rev. B **77**, 104435 (2008)
[15] P. Böni, D. Clemens, M. Senthil Kumar, C. Pappas, Physica B **267**, 320 (1999)
[16] U. Rücker, E. Kentzinger, chapter D4 in IFF Spring School "Probing the Nanoworld", FZ Jülich, series "Matter and Materials", Vol 34 (2007)
[17] S.K. Sinha, E.B. Sirota, S. Garoff, H.B. Stanley, Phys. Rev. B **38**, 2297 (1988)
[18] V. Holý, J. Kubena, I. Ohlídal, K. Lischka, W. Plotz, Phys. Rev. B **47**, 15896 (1993)
[19] V. Holý, T. Baumbach, Phys. Rev. B **49**, 10668 (1994)
[20] J. Stettner, Dissertation, Christian-Albrechts-Universität zu Kiel (1995)
[21] D.R. Lee, S.K. Sinha, D. Haskel, Y. Choi, J.C. Lang, S.A. Stepanov, G. Strajer, Phys. Rev. B **68** 224409 (2003)
[22] D.R. Lee, S.K. Sinha, C.S. Nelson, J.C. Lang, C.T. Venkataraman, G. Strajer, R.M. Osgood III, Phys. Rev. B **68** 224409 (2003)
[23] S. Disch, E. Wetterskog, R.P. Hermann, G. Salazar-Alvarez, P. Busch, Th. Brückel, L. Bergström, S. Kamali, Nano Lett. 11, 1651 (2011)
[24] E. Josten, E. Wetterskog, J.W. Andreasen, E. Brauweiler-Reuters et. al., to be published

D 3 Powder and Single Crystal Diffractometry: Chemical and Magnetic Structures

M. Meven
Institut für Kristallographie
RWTH Aachen

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Many mechanical, thermal, optical, electrical and magnetic properties of solid matter depend significantly on its atomic structure. Therefore, a good understanding of the physical properties needs not only the knowledge about the particles inside (atoms, ions, molecules) but also about their spatial arrangement. For most cases diffraction is *the* tool to answer questions about the atomic and/or magnetic structure of a system. Beyond this, neutron diffraction allows to answer questions where other techniques fail.

1.1 Crystallographic Basics

In the ideal case a complete solid matter consists of small identical units (same content, same size, same orientation like sugar pieces in a box). These units are called unit cells. A solid matter made of these cells is called a single crystal. The shape of a unit cell is equivalent to a parallelepiped that is defined by its base vectors $\mathbf{a_1}$, $\mathbf{a_2}$ und $\mathbf{a_3}$ and that can be described by its lattice constants *a*, *b, c*; α, β and γ (fig. 1). Typical lengths of the edges of such cells are between a few and a few ten Ångström (1 Å=10^{-10} m). The combination of various restrictions of the lattice constants between $a \neq b \neq c$; $\alpha \neq \beta \neq \gamma \neq 90°$ (triclinic) and $a = b = c$; $\alpha = \beta = \gamma = 90°$ (cubic) yields seven crystal systems. The request to choose the system with the highest symmetry to describe the crystal structure yields fourteen Bravais lattices, seven primitive and seven centered lattices.

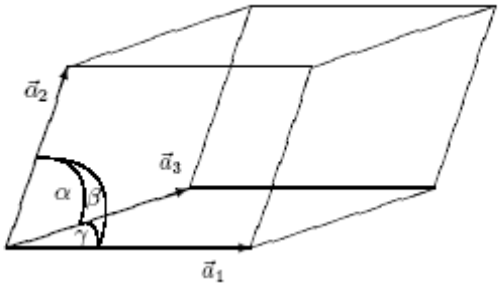


Fig. 1: *Unit cell with $|a_1|=a$, $|a_2|=b$, $|a_3|=c$, α, β, γ*

Each unit cell contains one or more particles *i*. The referring atomic positions $\mathbf{x_i}=x_i*\mathbf{a_1} + y_i*\mathbf{a_2} + z_i*\mathbf{a_3}$ are described in relative coordinates $0 \leq x_i; y_i; z_i < 1$. The application of different symmetry operations (mirrors, rotations, glide mirrors, screw axes) on the atoms in one cell yield the 230 different space groups (see [1]).

The description of a crystal using identical unit cells allows the representation as a three-dimensional lattice network. Each lattice point can be described as the lattice vector $\mathbf{t} = u*\mathbf{a_1} + v*\mathbf{a_2} + w*\mathbf{a_3}$; $u, v, w \in \mathbf{Z}$. From this picture we get the central word for diffraction in crystals; the *lattice plane* or *diffraction plane*. The orientations of these planes in the crystal are described by the so called *Miller indices h*, *k* and *l* with $h, k, l \in \mathbf{Z}$ (see pic. 2). The reciprocal base vectors $\mathbf{a*_1}$, $\mathbf{a*_2}$, $\mathbf{a*_3}$ create the reciprocal space with: $\mathbf{a*_i} * \mathbf{a_j} = \delta_{ij}$ with $\delta_{ij}=1$

for $i=j$ and $\delta_{ij}=0$ for $i \neq j$. Each point $\mathbf{Q}=h^*\mathbf{a}^*_1 + k^*\mathbf{a}^*_2 + l^*\mathbf{a}^*_3$ represents the normal vector of a (*hkl*) Plane. Each plane cuts the crystal lattice along its base vectors $\mathbf{a_1}$, $\mathbf{a_2}$ and $\mathbf{a_3}$ at $1/h^*\mathbf{a_1}$, $1/k^*\mathbf{a_2}$ and $1/l^*\mathbf{a_3}$. A Miller index of zero means that the referring axis will be cut in infinity. Thus, the lattice plane is parallel to this axis.

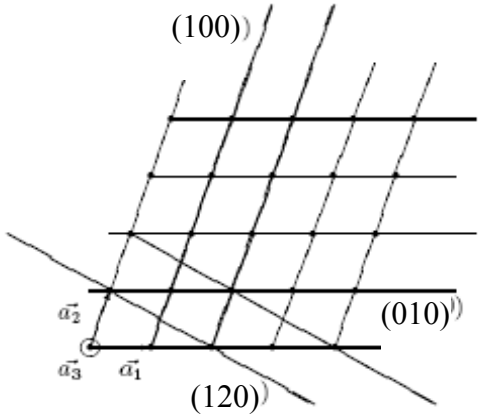


Fig. 2: *Different lattice planes in a crystal lattice, a_3 = viewing direction*

The atoms in a unit cell are not rigidly fixed at their positions. They oscillate around their positions (e.g. thermal excitation). A simple description for this is the model of coupled springs. In this model atoms are connected via springs whose forces describe the binding forces between the atoms (e.g. van der Waals, Coulomb, valence). The back driving forces of the springs are proportional to the deviation x_i of the atoms from their mean positions and to the force constant D, thus. $F = -D^*\Delta x$ (harmonic approximation).
Therefore, the atoms oscillate with $x_i = A_i^*\sin(\nu^* t)$ around their mean positions with the frequency ν and the amplitude A_i. Both, ν and A_i are influenced by the force constant D_j of the springs and the atomic masses m_i of the neighbouring atoms. The resulting lattice oscillations are called phonons in reference to the photons (light particles) in optics, which as well transport energy in dependence of their frequency. A more complex and detailed description of phonons in dependence on the lattice structure and the atomic interaction effects is given in lattice dynamics. In the harmonic approximation the displacements of an atom can be described with an oscillation ellipsoid. This ellipsoid describes the preferred volume in which the atom is placed. Its so called mean square displacements (MSD) U^i_{jk} represent the different sizes of the ellipsoid along the different main directions *j, k* in the crystal. The simplest case is a sphere with isotropic MSD B_i. In the next paragraph MSD are discussed from the point of view of diffraction analysis.
A full description of a single crystal contains information about lattice class, lattice constants and unit cell, space group and all atomic positions and their MSD. If the occupancy of one or more positions is not exactly 100%, e.g. for a mixed crystal or a crystal with deficiencies there has to be used also an occupancy factor.

1.2 Structure Determination with Diffraction

Diffraction means coherent elastic scattering of a wave on a crystal. Because of the quantum mechanical wave/particle dualism x-rays as well as neutron beams offer the requested wave properties:

Electrons: $E = h\nu$; $\lambda = c/\nu$
Neutrons: $E_{kin} = 1/2 * m_n * v^2 = h\nu = p^2/2m_n$; $\lambda = h/p$; $p \sim \sqrt{(m_n k_B T)}$

h: Planck's constant; ν: oscillation frequency; λ: wavelength; c: light speed; p: impact; m_n: neutron mass; k_B: Boltzmann constant; T: temperature

Only the scattering cross section partners are different (x-rays: scattering on the electron shell of the atoms; neutrons: core (and magnetic) scattering) as explained in detail below. In scattering experiments the information about structural properties is hidden in the scattering intensities I.
In the following pages we will discuss only elastic scattering ($\lambda_{in}=\lambda_{out}$). The scattering cross section of the radiation with the crystal lattice can be described as following:
Parallel waves of the incoming radiation with constant λ are reflected by lattice planes which are ordered parallel with a constant distance of d. This is very similar to a light beam reflected by a mirror. The angle of the diffracted beam is equal to the angle of the incoming beam, thus the total angle between incoming and outgoing beam is 2Θ (see fig. 3).

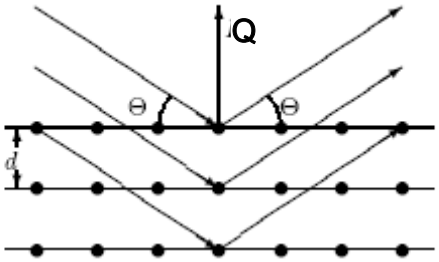


Fig. 3: *Scattering on lattice planes*

The overlap of all beams diffracted by a single lattice plane results in constructive interference only if the combination of the angle Θ, lattice plane distance d and wavelength λ meets Bragg's law:

$$2d \sin \Theta = \lambda$$

The largest distance $d_{hkl} = |\mathbf{Q}|$ of neighboured parallel lattice planes in a crystal is never larger than the largest lattice constant $d_{hkl} \leq \max(a; b; c)$. Therefore, it can only be a few Å or less. For a cubic unit cell ($a = b = c$; $\alpha = \beta = \gamma = 90°$) this means:
$d_{hkl} = a/\sqrt{(h^2+k^2+l^2)}$

With increasing scattering angle also the indices (*hkl*) increase while the lattice plane distances shrink with a lower limit of $d_{min} = \lambda/2$. Therefore, scattering experiments need wavelengths λ in the same order of magnitude of the lattice constants or below. This is equal to x-ray energies of about 10 keV or neutron energies about 25 meV (thermal neutrons).

Ewald Construction: In reciprocal space each Bragg reflex is represented by a point $\mathbf{Q} = h\mathbf{a}^*_1 + k\mathbf{a}^*_2 + l\mathbf{a}^*_3$. A scattered beam with the wave vector **k** fulfills Braggs law if the relationship $\mathbf{k} = \mathbf{k_0} + \mathbf{Q}$, $|\mathbf{k}|=|\mathbf{k_0}|=1/\lambda$ is true, as shown in fig. 4. During an experiment the available reciprocal space can be described by an Ewald sphere with a diameter of $2/\lambda$ and the (000)-point as cross point of $\mathbf{k_0}$ direction and the centre of the diameter of the sphere. The rotation of the crystal lattice during the diffraction experiment is equal to a synchronous movement of the reciprocal lattice around the (000)-point. If Bragg's law is fulfilled, one point (*h k l*) of the reciprocal lattices lies exactly on the Ewald sphere. The angle between the **k**-vector and the $\mathbf{k_0}$-vektor is 2Θ. The limited radius of $1/\lambda$ of the Ewald sphere limits also the visibility of (*h k l*) reflections to $|\mathbf{Q}| < 2/\lambda$.

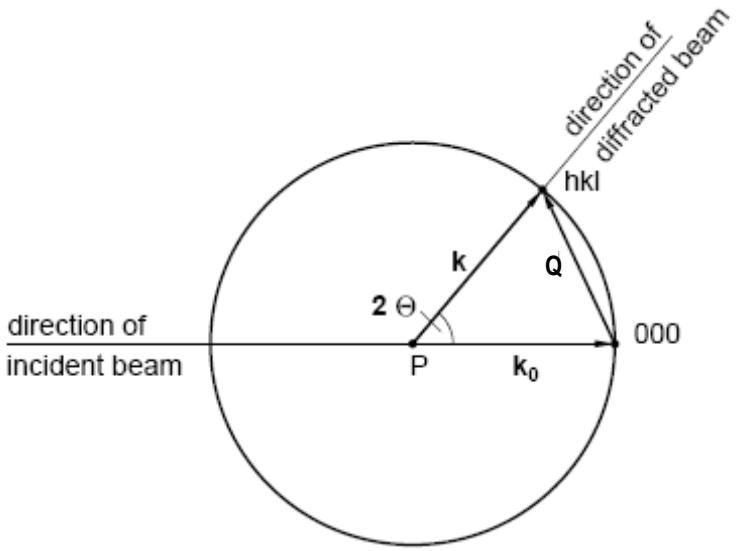


Fig. 4: *Ewald construction*

Determination of the unit cell: Following Braggs law the scattering angle 2Θ varies (for λ=const.) according to the lattice distance d_{hkl}. Thus for a given λ and known scattering angles 2Θ one can calculate the different *d* values of the different layers in the lattice of a crystal. With this knowledge is possible to determine the lattice system and the lattice constants of the unit cell (although not always unambigously!).

Atomic Positions in the unit cell: The outer shape of a unit cell does not tell anything about the atomic positions $\mathbf{x_i} = (x_i\, y_i\, z_i)$ of each atom in this cell. To determine the atomic positions one has to measure also the quantities of the different reflection intensities of a crystal. This works because of the relationship between the intensities of Bragg reflections and the specific

cross section of the selected radiation with each element in a unit cell. Generally one can use the following formula for the intensity of a Bragg reflection (*h k l*) with **Q** (kinetic scattering theory):

$$I_{hkl} \sim |F_{hkl}|^2 \text{ with } F_{hkl} = \sum^{n}_{i=1} s_i(\mathbf{Q}) \exp(2\pi\iota(hx_i+ky_i+lz_i))$$

The scattering factor *F* is a complex function describing the overlap of the scattering waves of each atom *i* (*n* per unit cell). $s_i(\mathbf{Q})$ describes the scattering strength of the *i*-th atom on its position $\mathbf{x}_i$ in dependence of the scattering vector **Q**, which depends on the character of cross section as described below.
In this context one remark concerning statistics: For measurements of radiation the statistical error σ is the square root of the number of measured events, e.g. x-ray or neutron particles. Thus, 100 events yield an error of 10% while 10,000 events yield an error of only 1%!

Mean Square Displacements (MSD): Thermal movements of atoms around their average positions reduce Bragg intensities in a diffraction experiment. In the simplest (=isotropic) description the parameter B_i is used to define a shell of electron or nucleus density around the average atomic position where the atom oscillates harmonically. The reduced probability to find an electron/nucleus at the average atomic position attenuates also the scattering probability. For higher temperatures (above a few Kelvin) the MSD B_i of the atoms increase linearly to the temperature *T*, this means $B \sim T$. Near a temperature of 0 K the MSD become constant with values larger than zero (zero point oscillation of the quantum mechanical harmonic oscillator). In the structure factor the true scattering capability s_i of the *i*-th atom has to be corrected by an angle-dependent factor (the so called Debye-Waller factor):

$$s_i(\mathbf{Q}) \rightarrow s_i(\mathbf{Q}) * \exp(-B_i(\sin\Theta_Q/\lambda)^2)$$

This Debye-Waller factor decreases with increasing temperature and yields an attenuation of the Bragg reflection intensities. At the same time this factor becomes significantly smaller with larger $\sin\Theta/\lambda \sim |\mathbf{Q}|$. Therefore, especially reflections with large indices become weaker. An improved description of probability density with anisotropic MSD U_{ij} contains the following exponential function:

$$s_i(\mathbf{Q}) \rightarrow s_i(\mathbf{Q}) * \exp(-2\pi^2(U^i_{11}\, h^2a^{*2} + U^i_{22}\, k^2b^{*2} + U^i_{33}\, l^2c^{*2} + \\ + 2U^i_{13}\, hl\, a^*c^* + 2U^i_{12}\, hk\, a^*b^* + 2U^i_{23} kl\, b^*c^*))$$

Here the U_{ij} describe the dimensions of an ellipsoid instead of a shell. The transformation between *B* and U_{eq} (isotropic MSD calculated from the anisotropic U_{ij} with identical volume) is: $B = 8\pi^2 U_{eq}$
For some structures the experimentally determined MSD are significantly larger than from the harmonic calculations of the thermal movement only. Static local deformations, point defects, mixed compounds, anharmonic oscillations or double well potentials (two energetically equal atomic positions very near to each other where an atom has a 50%/50% chance to occupy one position or the other) can cause this additional contribution to the pure thermal Debye-Waller factor. In the following text only the term MSD will be used to avoid misunderstandings.

1.3 Comparison of X-ray and Neutron Radiation

The different nature of interaction with matter of x-rays and neutrons explains why many studies use both techniques. The following table and picture show the main similarities and differences between the two radiation types.

Properties	X-ray/γ	Neutrons
Mass [kg]	0	$1.673*10^{-27}$
Energy [eV]	10^3 - 10^6	10^{-3} – 100 0.025 (thermal)
magn. Moment	no	Yes
Wave length λ [Å]	0.3 - 3 1.5 (Cu-Kα)	0.3 – 20 1.8 (thermal)
typ. speed [m/s]	$3*10^8$	2500 (thermal)
Interaction with	e- shell Z specific	cores/isotopes spin

Fig. 5: *Properties of X-rays vs. neutrons*

X-Ray Radiation interacts as electromagnetic radiation only with the electron density in a crystal. This contains all electrons whether they contribute to a chemical bond or not. The electronic scattering capability s – the so called atomic form factor $f(\sin\Theta/\lambda)$ or shorter f_Z - of an atom depends on the number Z of its shell electrons ($f(\sin(\Theta=0)/\lambda) = Z$). To be exact, $f(\sin\Theta/\lambda)$ is the Fourier transform of the radial electron density distribution $n_e(r)$: $f(\sin\Theta/\lambda) = \int_0^\infty 4\pi^2 n_e(r) \sin(\mu r)/\mu r \, dr$ with $\mu = 4\pi \sin\Theta/\lambda$. Heavy atoms with many electrons contribute much stronger to reflection intensities ($I \sim Z^2$) than light atoms with less electrons. The reason for the $\sin\Theta/\lambda$-dependence of f_Z is the diameter of the electron shell. It has the same order of magnitude as the wavelength λ and cannot be described as point like scattering centre. Thus, for large scattering angles the atomic form factors vanish and also the reflection intensities relying on them. The atomic form factors are derived from theoretical spherical electron density functions (e. g. Hartree-Fock). The resulting $f(\sin\Theta/\lambda)$ curves of all elements (separated for free atoms and ions) are listed in the international tables. Their analytical approximation is described by seven coefficients (c, a_i, b_i; $1 \leq i \leq 3$), see [1].

Neutron Radiation interacts with the cores and the magnetic moments of atoms. The analogon to the x-ray form factor (the scattering length b_c) is therefore not only dependent on the element but the isotope. At the same time b-values of elements neighboured in the periodic table can differ significantly. Nevertheless, the scattering lengths do not differ around several orders of magnitude like in the case of the atomic form factors f_Z . Therefore, in a compound with light and heavy atoms the heavy atoms do not dominate necessarily the Bragg intensities. Furthermore, the core potential with a diameter about 10^{-15}Å is a point like scattering centre and thus the scattering lengths b_c are independent from the Bragg angle and

sinΘ/λ respectively. This results in large intensities even at large scattering angles. The magnetic scattering lengths b_m can generate magnetic Bragg intensities comparable in their order of magnitude to the intensities of core scattering. On the other hand side the magnetic scattering lengths are strongly dependent on the sinΘ/λ value due to the large spacial distribution of magnetic fields in a crystal. Similarly to the x-ray form factors the b_m(sinΘ/λ) curves are listed in the International Tables. Therefore, it is easy to measure magnetic structures with neutrons and to separate them from the atomic structure.

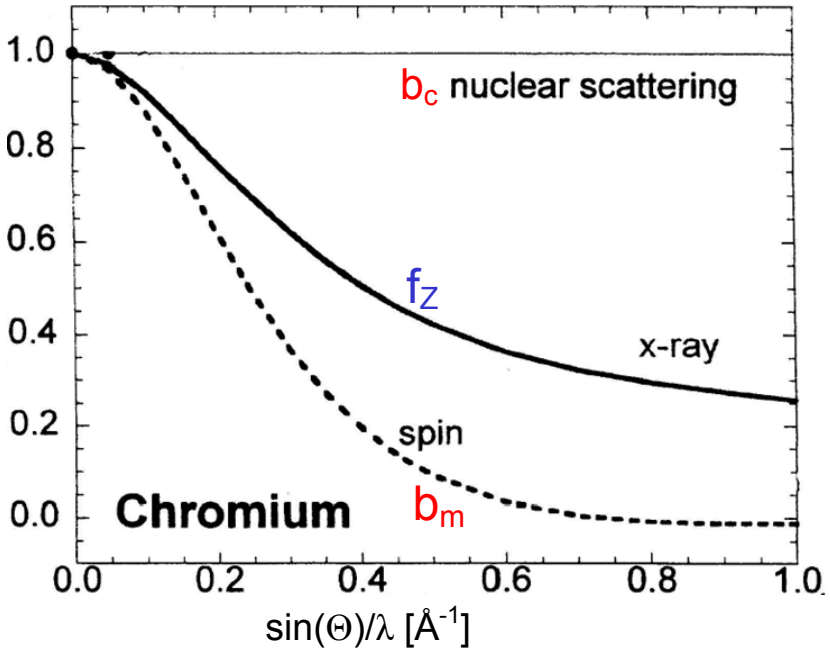


Fig. 6: *Normalized form factor f_Z and scattering lengths b_c and b_m for chromium*

Comparison: In summary in the same diffraction experiment the different character of x-ray and neutron radiation yield different pieces of information that can be combined. X-rays yield electron densities in a crystal while neutron scattering reveals the exact atomic positions. This fact is important because for polarised atoms the core position and the centre of gravity of electron densities are not identical any more. In compounds with light and heavy atoms structural changes driven by light elements need additional diffraction experiments with neutrons to reveal their influence and accurate atomic positions respectively. One has to take into account also that for x-rays the intensities depend in two ways on sinΘ/λ: Once by the atomic form factor f_Z, and twice by the temperature dependent Debye-Waller factor (see above). The first dependence vanishes if using neutron diffraction with b_c=const. and decouples the structure factors from the influence of the MSD. In general this yields much more accurate MSD U_{ij} especially for the light atoms and might be helpful to reveal double well potentials in (partially) disordered compounds.

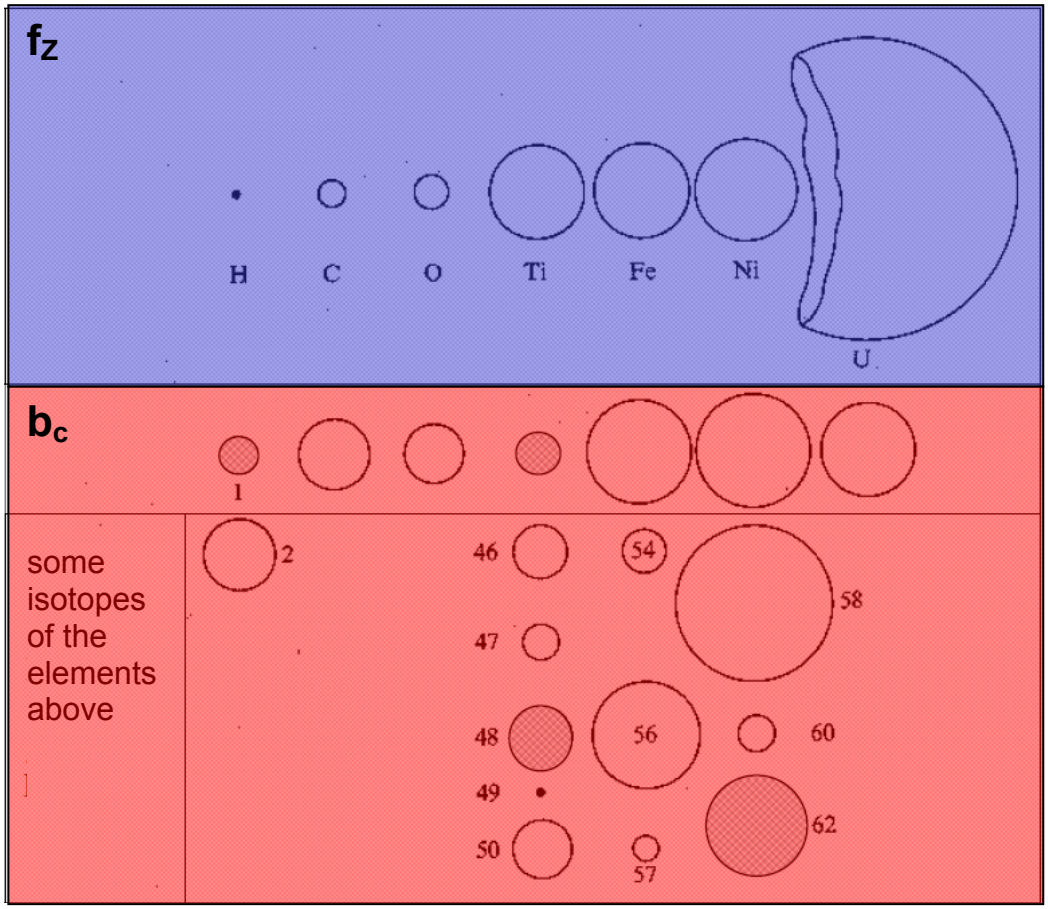


Fig. 7: *Comparison of f_Z and b_c for some elements and isotopes*
(from E.C. Bacon, Neutron Diffraction, Clarendon Press, 1975)

1.4 Special Effects and Aspects

From the relation $I \sim |F|^2$ one can derive that the scattering intensities of a homogenous illuminated sample increases with its volume. But there are other effects than MSD that can attenuate intensities. These effects can be absorption, extinction, polarization and the Lorentz factor:

Absorption can be described by the Lambert-Beer law:

$I = I_0 \exp(-\mu x)$, μ/cm^{-1} = linear absorption coefficient, x/cm = mean path through sample

The linear absorption coefficient is an isotropic property of matter and depends on the wavelength and kind of radiation. For x-rays penetration depths are only a few millimetre or below (e.g. for silicon with $\mu_{\mathrm{MoK}\alpha}$=1.546 mm^{-1}, $\mu_{\mathrm{CuK}\beta}$=14.84 mm^{-1} with penetration depths of 3 mm and 0.3 mm respectively). This limits transmission experiments to sample diameter of typically below 0.3 mm. To correct bias of intensities due to different scattering paths through

the sample one has to measure accurately the sample size in all directions. Even for sphere like samples the mean path lengths depend on 2Θ. In addition the sample environment must have an extraordinary small absorption.
Thermal neutrons have for most elements a penetration depth of several centimeters. Thus, sample diameters of several millimeters together with large and complex sample environments (furnaces, magnets, etc.) can be used. On the other hand side one needs sufficiently large samples for neutron diffraction which is often a delicate problem.

Extinction reduces also radiation intensities although its character is completely different form that of absorption. The principle of the *extinction effect* (not to be mixed up with crystallographic *extinction rules*!) can be explained quite easily by taking into account that each diffracted beam can be seen as a new primary beam for the neighbouring lattice planes. Therefore, the diffracted beam becomes partially backscattered towards the direction of the very first primary beam (Switch from kinetic to dynamic scattering theory!). Especially for very strong reflections this effect can reduce intensities dramatically (up to 50% and more). Condition for this effect is a merely perfect crystal.
Theoretical models which include a quantitative description of the extinction effect were developed from Zachariasen (1962) and Becker and Coppens [2, 3, 4, 5, 6]. These models base on an ideal spherical mosaic crystal with a very perfect single crystal (primary Extinction) or different mosaic blocks with almost perfect alignment (secondary Extinction) to describe the strength of the extinction effect. It is also possible to take into account anisotropic extinction effect if the crystal quality is also anisotropic. Nowadays most refinement programs [7] include extinction correction. In general extinction is a problem of sample quality and size and therefore more commonly a problem for neutron diffraction and not so often for x-ray diffraction with much smaller samples and larger absorption. Sometimes shorter wavelengths where extinction effects become weaker can be used as solution.

Multiple scattering occurs if the diffracted beam of the first lattice plane ($h_1k_1l_1$) works as primary beam for a second non equal lattice plane ($h_2k_2l_2$) that by accident also fulfils Bragg's law. The result is a diffracted beam virtually generated by a third lattice plane ($h_3k_3l_3$). If the structure factors of the two first planes, F_{h1k1l1}, F_{h2k2l2} are strong, the measured intensity I_{h3k3l3} might be larger than the true one generated from F_{h3k3l3}.

Polarisation*:* X-ray radiation is electromagnetic radiation. Therefore, the primary beam of an x-ray tube is not polarized. The radiation hits the sample under an diffraction angle of □ where it can be separated into two waves of same intensity, firstly with an electrical field vector parallel $E_{||}$ and secondly perpendicular $E_{\perp}$ towards the □-axis. Whilst the radiation with $E_{||}$ will not be attenuated the radiation with $E_{\perp}$ undergoes an attenuation with $E_{\perp} \rightarrow \cos(2\Theta)\, E_{\perp}$. The polarization factor P for the attenuation has then the following formula ($I \sim E^2$):

$$P = (1+\cos(2\Theta)^2)/2$$

Additional optical components like monochromator crystals also have an impact on the polarization and have to be taken into account accordingly.

Lorentz factor: The Lorentz factor L is a purely geometrical factor. It results from the effectively elongated stay of the sample near the reflection position during an ω- and ω/2Θ-scan respectively towards higher 2Θ values for the same angular speed $\Delta\omega / \Delta t$:

$L = 1/\sin(2\,\Theta)$

This has to be taken into account for any kind of radiation in an diffraction experiment.

Technical limits: The different interactions of x-ray and neutron radiation with the atoms in a crystal make neutrons apparently the better choice for diffraction experiments which focus on atomic positions, mean square displacements and magnetism. X-rays are preferable for studies on electron densities and chemical bonds. But one has to take into account the available flux of x-rays and neutrons respectively. Diffraction with a monochromatic beam needs a sharp band of energies/wavelengths in the order of $\Delta\lambda/\lambda<10^{-2}$ - 10^{-3} or even smaller. For such a small bandwidth the flux of neutrons is several orders of magnitude smaller than the flux of x-rays of a corresponding synchrotron source or x-ray tube in the laboratory. The reason for this is that in an x-ray tube most x-rays are generated in a small energy band, the characteristic lines of the tube target (K_α, K_β, etc.). Additional metal foils used as filter allow to cut off unwanted characteristic lines which yields quasi monochromatic radiation of a single wavelength. Neutrons generated by fission in a research reactor distribute to a broad spectrum of wavelengths. To reduce the bandwith one has to use a monochromator crystal. This reduces significantly the number of available neutrons for the diffraction experiment. Thus, the weak flux of neutrons and the weak cross section of neutrons with matter have to be compensated with large sample sizes of several millimeters. For the same reason the monochromatization of the neutrons is normally chosen to be not too sharp (resolution about $\Delta\lambda/\lambda\approx10^{-2}$ for neutrons, $\Delta\lambda/\lambda\approx10^{-4} – 10^{-5}$ for synchrotron).

1.5 From Measurement to Model

Goal of powder diffraction and single crystal diffraction is to compare experimental diffraction data with several structure models and to find the one that fits best. The differences between both methods are explained in the following sections. For instance, powder diffraction experiments yield I/2Θ diagrams while single crystal diffraction experiments yield a list of corresponding intensities and error bars of a set of Bragg reflections and therefore slightly different refinement procedures are used to find a structure solution. Nevertheless the general approach described for single crystal data analysis in this section is very similar to the one for powder diffraction which will not be discussed here.

To get a structural model from the experimentally collected integral Bragg intensities one needs several steps in advance. Firstly one has to make sure that all reflections are measured

properly (no shading, no $\lambda/2$-contamination, no multiple scattering). Damaged reflections have to be excluded from further treatment.
During data refinement not only the quantities of the relative intensities but also their errors are taken into account. The total statistical error □ of an integral intensity I_{obs} of a single reflection is calculated as following:

$$\sigma^2 = I_{obs} + I_{background} + (k\, I_{total})^2$$

The part $\sigma_0^2 = I_{total}$, $I_{total} = I_{obs} + I_{background}$ refers to the error caused by counting statistics. It contains as well the effective intensity I_{obs} as the contribution of the background. But other effects also influence the reproducibility of a measurement (and thus the total error), e.g. specific the instrumental errors. Those errors are collected in the so called *McCandlish-Factor* k and increase to the total error. Obviously, the total error cannot drop below the instrumental limit of the experiment and thus the impact of strong reflections does not become exaggerated in the refinement. The determination of k is done by measuring the same set of reflections (so called standard reflections) several times during a data collection. The mean variation of the averaged value represents k. In addition, the repeated measurement of standard reflections offers the opportunity to notice unwanted changes during experiment like structural changes or release from the sample holder.
To make sure the comparability of all reflections with each other, all intensities and errors are normalized to the same time of measurement (or monitor count rate) and undergo the Lorentz and (in the x-ray case) polarization correction.
Finally in advance of the data refinement there can be done a numerical (e.g. with DataP, [8]) or an empirical absorption correction if necessary. The quality of a measurement is checked in advance of the data refinement by comparing symmetry equivalent reflections and systematic extinctions to confirm the Laue group and space group symmetry. The result is written as internal R-value:

$$R_{int} = (\sum_{k=1}^{m}(\sum_{j=1}^{n_k}(<I_k> - I_j)^2))/(\sum_{k=1}\sum_{j=1}^{n_k}(I_j^2)_k)$$

R_{int} represents the mean error of a single reflection j of a group k of n_k symmetry equivalent reflections, corresponding to its group and the total number m of all symmetrically independent groups. Therefore R_{int} is also a good mark to check the absorption correction. After these preliminary steps one can start the final data refinement.
At the beginning one has to develop a structural model. The problem with that is that we measure only the absolute values $|F_{hkl}|$ and not the complete structure factor $F_{hkl} = |F_{hkl}|\exp(\imath\,□)$ including its phase □. Therefore, generally the direct Fourier transform of the reflection information F_{hkl} from reciprocal space into the density information ρ in the direct space (electron density for x-rays, probability density of atomic cores for neutrons) with

$$\rho(\mathbf{x}) \sim \sum_h\sum_k\sum_l F_{hkl}\exp(-2\pi(hx+ky+lz))$$

is not possible. This can be done only by direct methods like Patterson, heavy atom method or anomal dispersion for x-rays.

In the so called refinement program a given structural model (space group, lattice constants, atomic form factors, MSD, etc.) are compared with the experimental data and fitted. In a least squares routine those programs try to optimize (typically over several cycles) the free parameters to reduce the difference between the calculated structure factors F_{calc} and intensities $|F_{calc}|^2$ respectively and the experimentally found F_{obs} and $|F_{\mathrm{obs}}|^2$ respectively. To quantisize the quality of measurement there are several values in use:

1. *unweighted R*-value: $R_{\mathrm{u}} = \sum_{hkl} |F_{\mathrm{obs}}^2 - F_{\mathrm{calc}}^2| / \sum_{hkl} F_{\mathrm{obs}}^2$
This value gives the alignment of the whole number of reflections without their specific errors.

2. weighted *R*-value: $R_{\mathrm{w}} = (\sum_{hkl} w\,(F_{\mathrm{obs}}^2 - F_{\mathrm{calc}}^2)^2) / \sum_{hkl} w\,F_{\mathrm{obs}}^4$
This value represents the alignment of the whole number of reflections including their specific errors or weights ($w \sim 1/\sigma^2$). Sometimes weights are adopted in a way to suppress unwanted influence of the refinement algorithm by weak or badly defined reflections. Be aware that such corrections have to be done extremely carefully because otherwise the refinement adopts the data to the selected structural model and not the model to the experimental data!

3. *Goodness of Fit S*: $S^2 = (\sum_{hkl} w\,(F_{\mathrm{obs}}^2 - F_{\mathrm{calc}}^2) / (n_{\mathrm{hkl\text{-}reflections}} - n_{\mathrm{free\ parameter}})$

S should have a value near one if the weighting scheme and the structure model fit to the experimental data set.

2 Powder Diffraction

2.1 Method

Crystalline powder consists of a large number of randomly oriented microscopic single crystals, so called crystallites. Using a monochromatic beam a powder sample in a small capillary or on a flat table generates cones of diffracted radiation, the so called Debye Scherrer cones. Following Bragg's law the discrete angles at which intensity can be found by a photo film, single or area detector one can calculate the crystallographic unit cell and its fundamental symmetry (trigonal, monoclinic, hexagonal, orthorhombic, tetragonal, cubic). The projection of the integrated intensities from the detector area yields a diffractogram that presents the intensity distribution against 2Θ and Q respectively. In a Rietveld procedure the full reflection pattern (profiles and intensities) is analysed.

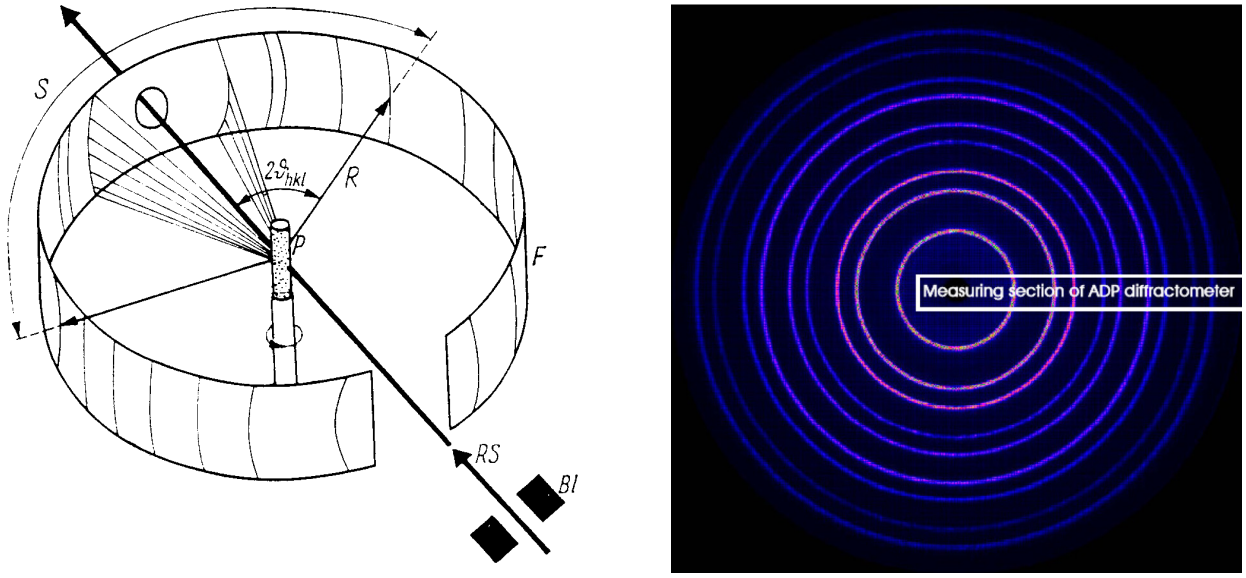


Fig. 8: *Powder diffraction principle and the Debye Scherrer cones on an image plate*

The main advantage of powder diffraction is the speed of data collections. At high flux sources a complete diffractogram can be collected within minutes. This allows not only to perform data analysis routinely and for many samples in a short time but can also be used to do in situ experiments where for instance a chemical reaction can be visualised by changes in a sequence of diffractograms measured one after each other. The other advantage with powder samples is the fact, that it is much easier to prepare a sample with a huge number of small crystals instead of one large single crystal. The main disadvantage of powder diffraction is that lattice planes that are crystallographically different will occur at the same position in the diffractogram due to the projection of diffraction intensities towards the diffraction angle. One problem of sample preparation can occur from crystallites of anisotropic shape (needles, plates). In powder samples these crystallites might order not fully randomly. As a result, the

Debye Scherrer cones are not homogeneously illuminated which yields diffractograms with misleading intensity distributions. This effect is called texture.

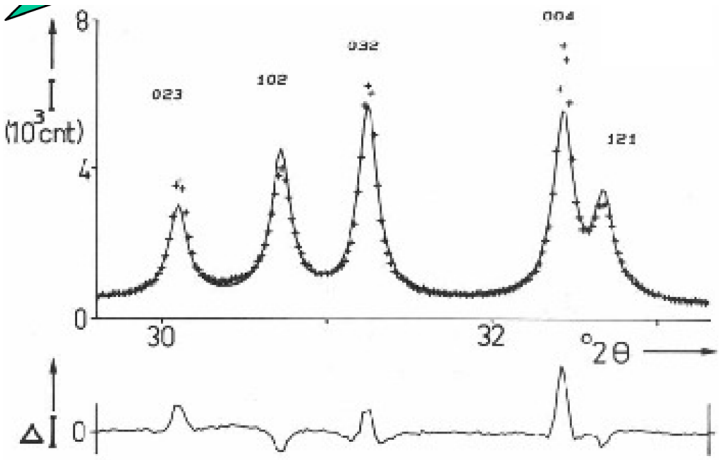
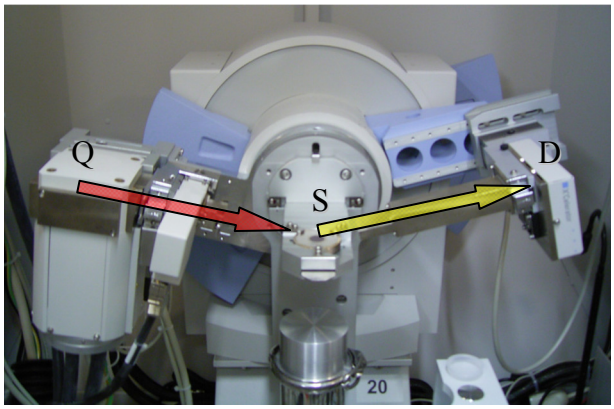


Fig. 9: *Diffractogram with experimental data (dots), model fit (line) and difference plot*

2.2 Instrumentation

Laboratory x-ray powder diffractometers use characteristic K_α radiation from a Cu, Mo, Ag or W targets in an x-ray tube. These instruments have a high flux and good resolution. Powder diffractometers at a synchrotron source offer additionally variable wavelengths, very high photon flux up to 10^{12} photons/s and very high resolution.
The following figure shows on the left a laboratory x-ray diffractometer. Its beam tube generates mainly characteristic K_α and K_β radiation while the amount of "Bremssrahlung" is rather small.
As the K_β radiation can be cut out instantly by a metal foil filter (e.g. Zn for Cu radiation) behind the beam tube no additional measures have to be taken to use a monochromatic primary beam. In this case the sample stage stays fixed at its position. The source moves clockwise while the detector does the same counter clockwise yielding at any time an angle between primary beam and sample of Θ and 2Θ between primary and secondary beam to fulfil Bragg's law.
An example for a thermal neutron powder diffractometer with fixed wavelengths of 1.11 Å, 1.55 Å or 2.54 Å is the instrument SPODI in the experimental hall of the neutron research source Heinz Maier-Leibnitz (FRM II) on the right. As the neutron source generates a broad spectrum of thermal neutrons this "white beam" needs to be monochromatised by a monochromator crystal first (e.g. Cu-220, Ge-311, etc.). The resulting monochromatic neutron flux is several orders of magnitude smaller than the photon flux at a synchrotron.
To overcome this problem the detector consists of a set of ^{3}He detectors covering about 160° in 2Θ and counting simultaneously to measure several diffractograms per day.



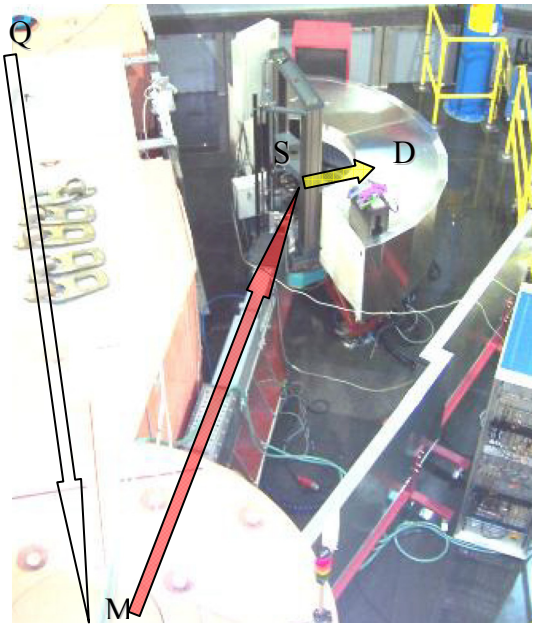


Fig. 10: *Up: Laboratory x-ray diffractometer with Θ/2Θ-Geometry (Bragg-Brentano)*

Right: Neutron powder diffractometer SPODI (Q: source, S: sample, D: detector, M: monochromator)

2.3 Examples

The method of powder diffraction warranties that no Braggs reflections inside the observed 2Θ range are missed. This is highly advantageous if samples of uncertain composition and/or structure have to be studied and can also be used to observe multiple phases simultaneously or behind each other, e.g. in temperature dependent experiments. As data collection in powder diffraction is rather fast these in situ experiments allow not only to study the results of a structural or chemical change but also the intermediate states.

In Situ Experiments on Zirconia Zirconia play an important role in many industrial applications like electrolytes, ion exchangers, catalysators and piezo-electrics. Therefore, the detailed understanding of its structural properties and synthesis are very important. A typical synthesis route is shown in fig. 11.

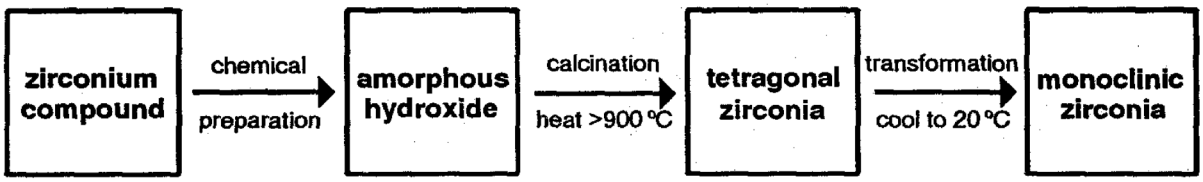


Fig. 11: *Wet chemical/furnacing route [9]*

Without going into details the following key aspects can be shown: The zirconium compound at the beginning of synthesis – $Zr(OH)_2SO_4 \cdot 3H_2O$ contains a huge amount of water/OH groups and shows amorphous behaviour. During the heating process a loss of water/OH groups is generated that results in a crystallisation in the tetragonal zirconia phase. After cooling down the structural phase transition from the tetragonal to monoclinic phase occurs during the cooling cycle but depends strongly on the dwelling time (for details see [9] and references within).

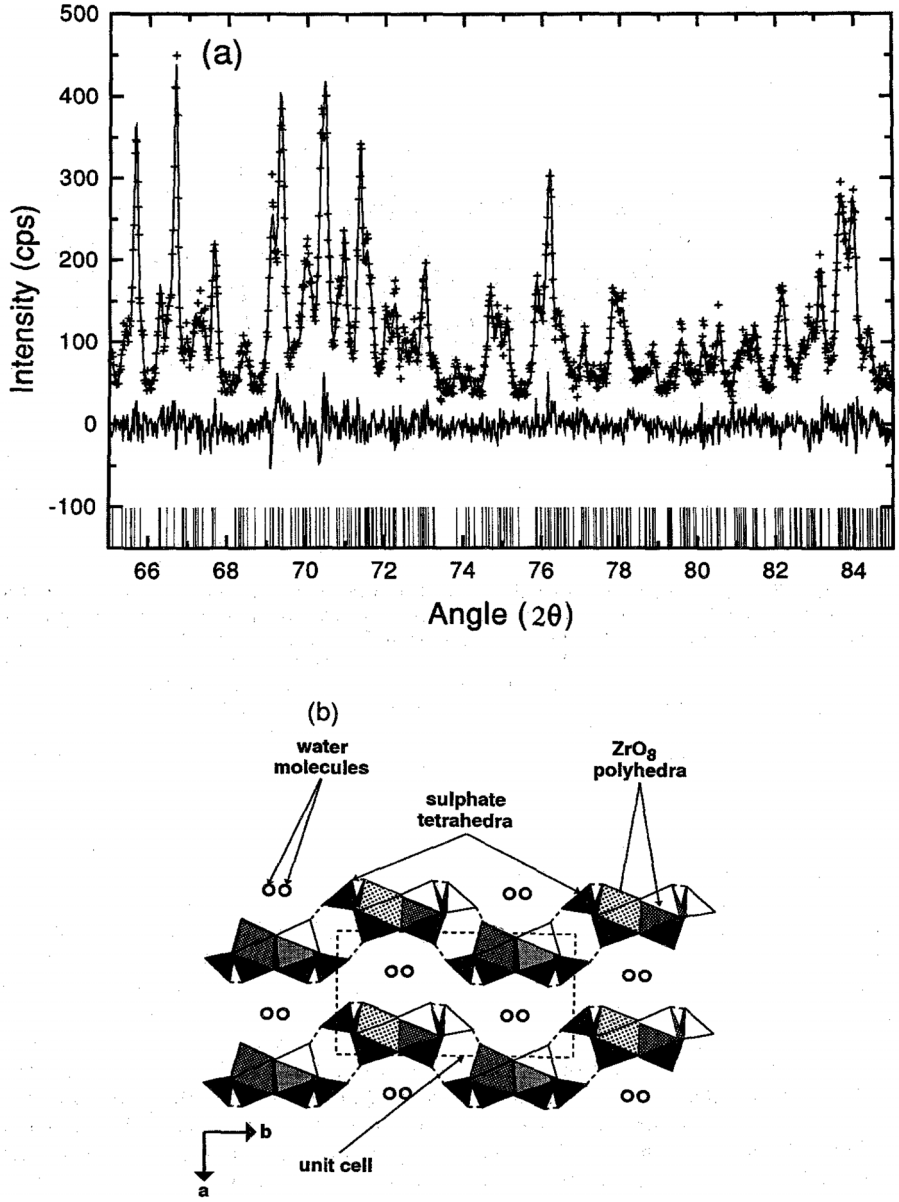


Fig. 12: *T dependent laboratory powder diffractogram of Zirconia [9]*
(a) Part (65-85°C) of the laboratory x-ray powder diffraction pattern obtained on $Zr(OH)_2SO_4 \cdot 3H_2O$ is shown to illustrate the detail and high quality that can be achieved using a fully monochromatized laboratory x-ray source. The data are displayed as a "Rietveld plot" in which the upper trace shows

*the observed data as dots and the calculated pattern as a solid line; the middle trace gives the difference between the observed and calculated data; and the lower set of vertical bars identifies each reflection. (b) A representation of the $Zr(OH)_2$ SO_4 $\cdot 3H_2O$ structure (after Gascoigne et al. 1994) illustrating zig-zag chains, viewed end on down the c-direction, which are hydrogen bonded together (broken lines) to form curraged sheets, parallel to **a** and **b**, which themselves are hydrogen bonded together via inter-layer water molecules [9].*

The figure above shows a part of a laboratory powder diffractogram and the corresponding Rietveld plot (perpendicular dashes below). The second half of this figure shows the tetragonal structure. The following figure shows the corresponding angular dispersive powder diffractograms for laboratory and synchrotron sources. The superior resolution of the synchrotron source is obvious.

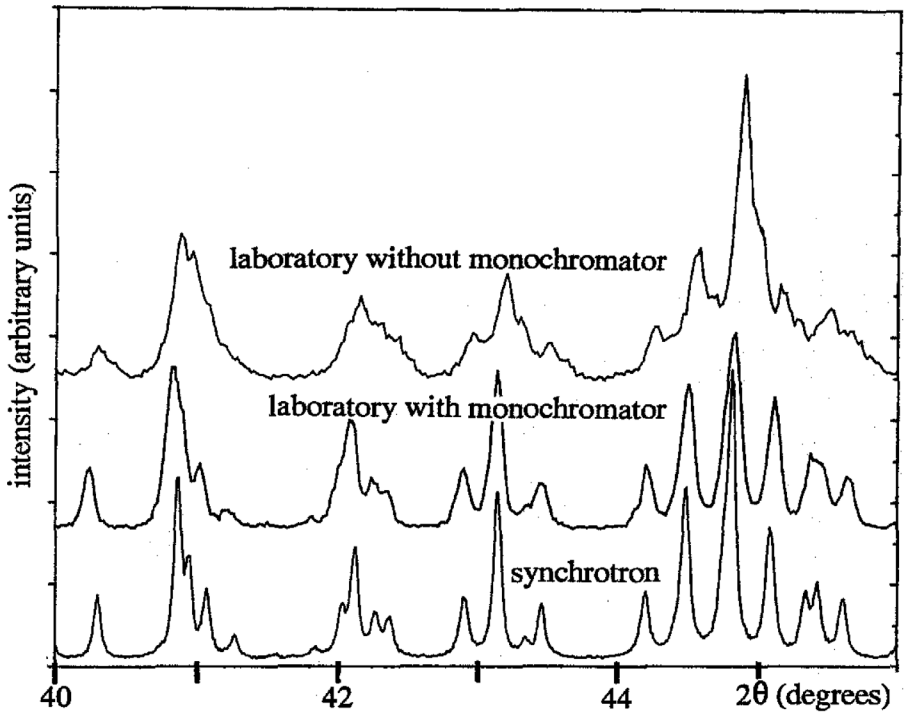


Fig. 13: *Comparison of laboratory and synchrotron powder diffractograms [9]*
Portions (considerably expanded, between 40° and 46°) of powder diffraction patterns for $Zr(OH)_2$ SO_4 $\cdot 3H_2O$ obtained using three different collection modes: the upper trace is with a conventional (Bragg-Brentano) laboratory x-ray powder diffractometer utilizing the full $K_{\alpha 1 \alpha 2}$-doublet from a copper x-ray source (mean wavelength λ=1.5418 Å); the middle trace is again with a laboratory x-ray source but with the $K_{\alpha 2}$ radiation removed by means of a germanium (100) monochromator [resultant $\lambda(K_{\alpha 1})$ = 1.5406 Å]; the lower trace is from a synchrotron (station 2.3 on the Daresbury source) powder diffractometer monochromated to λ = 1.51603 Å. the progressive improvements in pattern quality are self-evident, particularly with the closely-spaced multiple peaks around 41° 42° and 45°.

The final figure of this section presents the T dependent synchrotron energy dispersive powder diffractogram (EDD) which shows clearly the evolution and devolution of Bragg peaks during the different stages covering the range from the amorphous pattern of the hydroxide to the monoclinic oxide (The patterns were collected on station 9.7 of the

Daresbury synchrotron source using a diffraction angle of 2Θ=7.8° and a collection of 60s per pattern [9]).

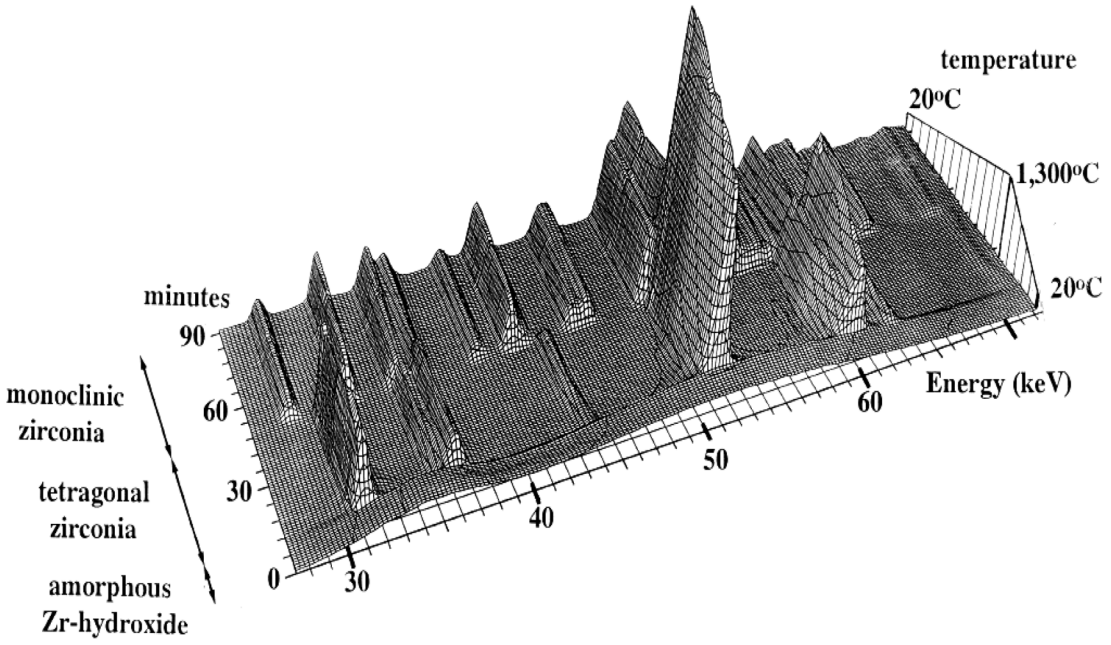


Fig. 14: *T dependent synchrotron powder diffractogram of Zirconia [9]*

3 Single Crystal Diffraction

3.1 Method

Classical angle dispersive single crystal diffractometers use a fixed constant wavelength λ. As the sample is one large single crystal, a lattice plane fulfilling Bragg's law will generate not a Debye-Scherrer cone but a single diffracted beam. The schematic drawing in fig. 15 shows a single crystal (black dot) in the centre of a Eulerian cradle with three axes ω, χ, φ. These three axes allow the crystallographic coordinate system of the sample to be oriented relatively to the coordinate system of the diffractometer towards any direction. Therefore any lattice plane can be oriented in a way that its diffracted beam hits the detector in the diffraction plane defined by the 2Θ axis. By measuring each lattice plane separately the problem of reflection overlaps as described above for powder diffraction can be overcome. Also certain anisotropic effects can be detected that stay hidden in powder diffractograms. But the not simultaneous measurement of Bragg intensities is very time consuming. This problem can be healed partially by using image plate detectors.

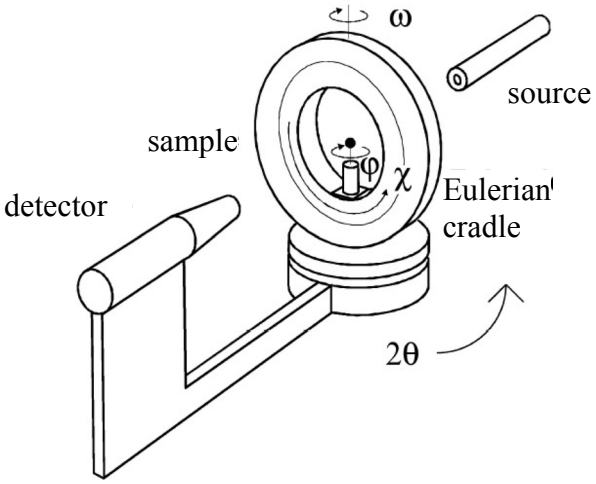


Fig. 15: *Scheme of a classical single crystal diffractometer with single detector and Eulerian cradle*

3.2 Instrumentation

An example for a classical neutron single crystal diffractometer is HEiDi from JCNS at the hot source at FRM II (fig. 16). The instrument offers discrete wavelengths between 1.2 Å and 0.4 Å and a large Q range.

The use of single detectors makes diffraction of single crystals a very time consuming task. Especially for x-ray diffraction analysis the development of new large area detectors, so called image plates with a laser read out at the end of last century helped to drop the necessary beam time for an experiment by up to one order of magnitude [10]. Comparable to the rotating crystal technique the sample is only rotated around one fixed axis φ. Special software transforms the collected Bragg intensities into an image of the intensity peak distribution in the reciprocal space.

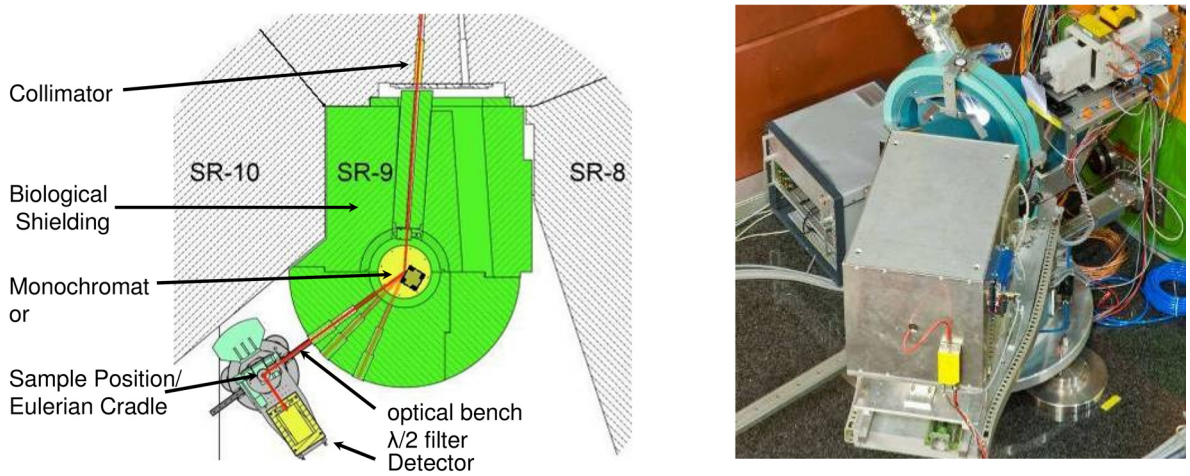


Fig. 16: *Classical neutron diffactometer HEiDi with Eulerian cradle and mounted cryostat*

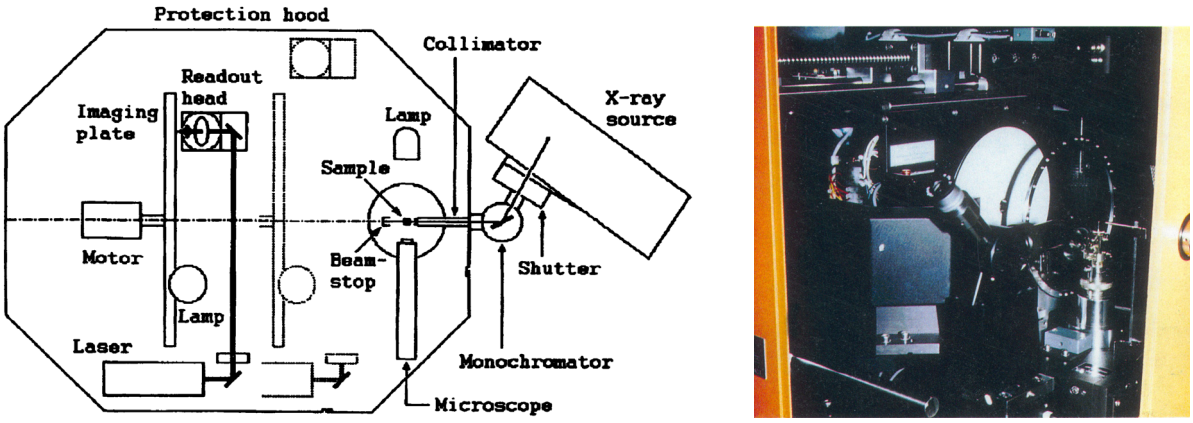


Fig. 17: *X-ray single crystal image plate diffractometer (IPDS STOE [10])*

3.3 Examples

Mixed Crystals - $ZrAs_xSe_y$: This metallic compound shows an anomaly in its electric resistivity below 16 Kelvin. The chemical structure of this compound at room temperature is tetragonal (P4/nmm, a = 3.78 Å, c = 8.14 Å). To investigate the question, whether the electric resistivity is related to a possible order/disorder or redistribution of the As and Se atoms on crystallographic *2a* and/or *2c* positions taking into account the possibility of vacancies requires the combination of x-ray and neutron diffraction [11].

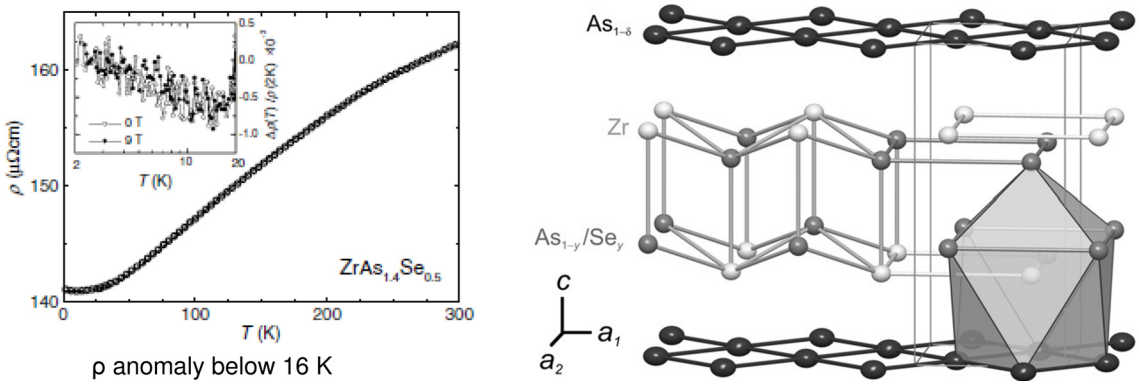


Fig. 18: *Electrical anomaly and As/Se distribution in $ZrAs_xSe_y$ [11]*

X-ray diffraction cannot distinguish between As and Se because of Z_{As}=33, Z_{Se}=34 are too similar. Nevertheless, x-ray diffraction can be used to determine the number of vacancies at the possible lattice positions of both elements. Chemical analysis is used to check the stochiometry of the samples. These constraints were used in a single crystal diffraction neutron experiment to determine the As and Se distribution in this compound taking into account the different neutron scattering lengths of As and Se with b_{As}=6.58 fm, b_{Se}=7.97 fm. The data shown in the table below were taken at three different temperatures, two above and

one below the anomaly. Please notice that at room temperature a Bragg data set was taken up to Q=0.78 Å^{-1} while the two low temperature Bragg data sets were taken up to Q=1.10 Å^{-1}. The increase in Q range by about 40% increases not only the number of available reflections by a factor of about 4 but drops the error bars of the atomic parameters by one order of magnitude!

The atomic positions of the elements in the unit cell of this compound are not randomly chosen. According to the space group of the unit cell of this compound there are symmetry restrictions. For instance, the first As position (3/4 1/4 0) stays fixed during the refinement process. Only the z positions of the other atoms and the anisotropic mean square displacements and occupancies of two of the three atomic positions can be set as free parameters. In combination with two additional parameters (scaling parameter for the normalisation of the sample volume and extinction parameter for eliminating extinction effects) the maximum number of free parameters is limited to 12.

Specimen		Crystal I		Crystal II[a]	
Chemical composition (WDXS)		$Zr_{1.000(3)}As_{1.595(3)}Se_{0.393(1)}$		$Zr_{1.000(2)}As_{1.420(3)}Se_{0.560(2)}$	
Crystal dimensions/mm		2 × 2 × 2		1.5 × 1.5 × 1.5	
Temperature/K		300		25.0(1)	2.3(1)
Radiation		Neutron		Neutron	
Diffractometer		RESI		HEiDi	
Wavelength/pm		104.9(1)		55.2(1)	
Space group		P4/*nmm*		P4/*nmm*	
Z		2		2	
Cell dimensions	*a*	378.30(7)		375.0(3)	
Neutron/pm	*c*	813.6(1)		803.7(1)	
$V/10^6$ pm^3		116.44		113.02	
Cell dimensions	*a*	375.76(2)		375.86(2)	
X-ray/pm	*c*	807.80(5)		805.45(7)	
$2\theta_{max}$/deg		110.8		75.0	
hkl range		±4	−6–+8		−6–+7
		−4–+3	−3–+8		−3–+7
		±12	−17–+16		±16
No. reflections		355	1679		1518
No. parameter		12	12		12
GOF		1.26	1.09		1.14
R1/*wR2*		0.051/0.132	0.027/0.057		0.023/0.047
Largest difference peak		0.74	0.97		0.96
Refined composition		$ZrAs_{1.48(10)}Se_{0.48}$	$ZrAs_{1.36(2)}Se_{0.62}$		$ZrAs_{1.33(2)}Se_{0.65}$

Fig. 19: *Table of Bragg data statistics with neutron single crystal diffraction [11]*

Atom	Site	*x*	*y*	*z*	U_{11}	U_{33}	U_{eq}	occ.
Zr	2*c*	1/4	1/4	0.2666(2)	0.0062(6)	0.0072(7)	0.0065(5)	1
				0.26505(3)	0.00247(7)	**0.00181(7)**	**0.00225(6)**	
				0.26506(2)	*0.00237(7)*	*0.00176(7)*	*0.00217(6)*	
As	2*a*	3/4	1/4	0	0.0089(8)	0.0072(8)	*0.0083(7)*	0.96(2)
					0.00671(9)	**0.00204(8)**	**0.00515(7)**	**0.979(4)**
					0.00644(8)	*0.00193(8)*	*0.00494(7)*	*0.975(4)*
As/Se	2*c*	1/4	1/4	0.6205(2)	0.0058(6)	0.0077(7)	*0.0065(5)*	0.52(8)/0.48
				0.62157(2)	**0.00153(6)**	**0.00233(6)**	**0.00180(5)**	**0.38(2)/0.62**
				0.62152(2)	*0.00162(6)*	*0.00238(6)*	*0.00187(5)*	*0.35(2)/0.65*

3 Data sets: RT (RESI, $ZrAs_{1.595(3)}Se_{0.393(1)}$), 25K and 2.3K (both HEiDi, $ZrAs_{1.420(3)}Se_{0.560(2)}$)

Fig. 20: *Results of the refinements [11]*

The high accuracy of the data concerning atomic positions, mean square displacements and occupancies and the combination of x-ray and neutron diffraction show unambigously that there are

- only As and vacancies on the 2a position,

- no vacancies but As and Se on the 2c position,
- no incalation of Zr in interstitial positions and
- no structural change down to 2.5 Kelvin.

From this result, a structural change can be excluded as cause for the observed anomaly of the electric resistivity.

HT Superconductors: $La_{2-x}Sr_xCuO_4$ is one of the cuprate superconductors with K_2NiF_4-structure (layered perovskite) for whose discovery the Noble prize was granted in 1988 (Bednorz and Müller [12]). Pure La_2CuO_4 is an isolator while doping with earth alcali metals (Ca^{2+}, Sr^{2+}, Ba^{2+}) on the La^{3+} lattice positions generates depending on the degree of doping superconductivity with a maximum T_c of 38 K for Sr doping of x=0.15 (see fig. 21).
Pure La_2CuO_4 undergoes at $T_{t\text{-}o}$=530 K a structural phase transition from the tetragonal high temperature phase (HTT) $F4/mmm$ (a=b=5.384 Å, c=13.204 Å, α=β=γ=90° at T=540 K) to the orthorhombic low temperature phase (LTO) Abma (a=5.409 Å, b=5.357 Å, c=13.144 Å, α=β=γ=90° at room temperature). For $La_{2-x}Sr_xCuO_4$ the phase transition temperature $T_{t\text{-}o}$ drops with increased doping and disappears above x=0.2.

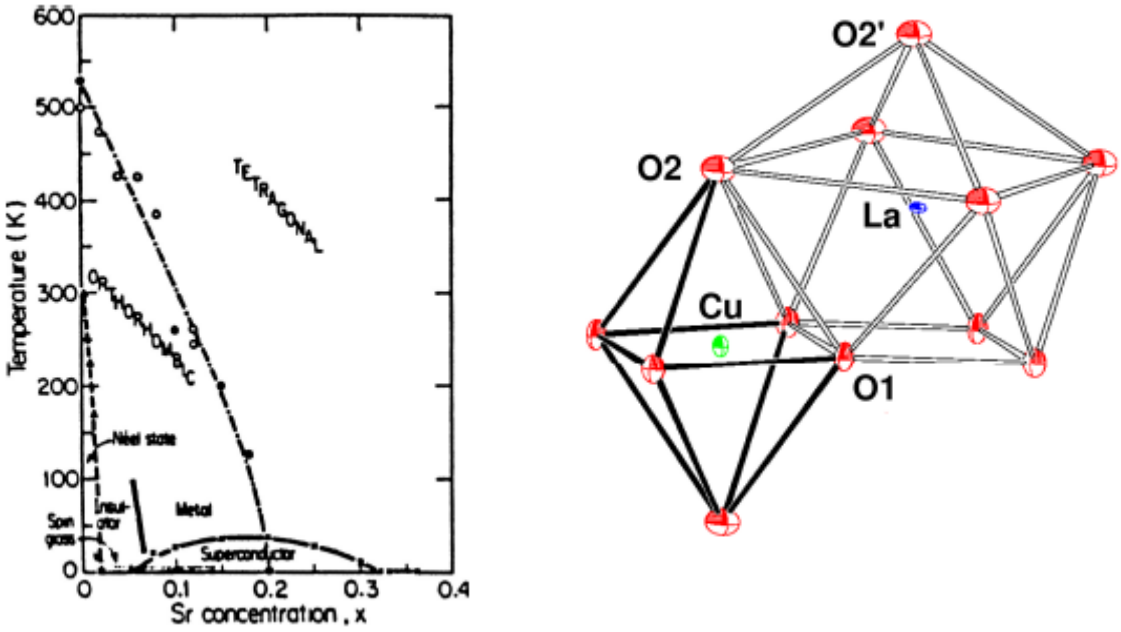


Fig. 21: *Phase diagram [13] and structural units, e.g. CuO_6 octahedra and LaO_9 polihedra*

The following aspects of structure analysis can be learned from this example:
Twinning: The structural phase transition yields a symmetry reduction that makes the single crystal separate into domains of identical structure with different but well defined orientations to each other. This effect is called *twinning*. As can be seen in fig. 22 the transition into the low temperature phase tilts the CuO_6 octahedrons around their [010] axis. The two axes of identical length in the HTT phase, a_1 and a_2, are not equal in the LTO phase anymore. Instead, the longer one becomes the new a axis, the shorter one becomes the b axis. Whether a_1 or a_2 becomes the new a axis depends only on the real structure of the crystal, for instance

grain boundaries or point defects. Two equivalent crystallographic space groups describe the LTO phase:

A*bma* ($a_1 \rightarrow a$, $a_2 \rightarrow b$) and B*mab* ($a_1 \rightarrow b$, $a_2 \rightarrow a$)

For the structure factors in the LTO is valid: $F_{Abma}(hkl)=F_{Bmab}(khl)$

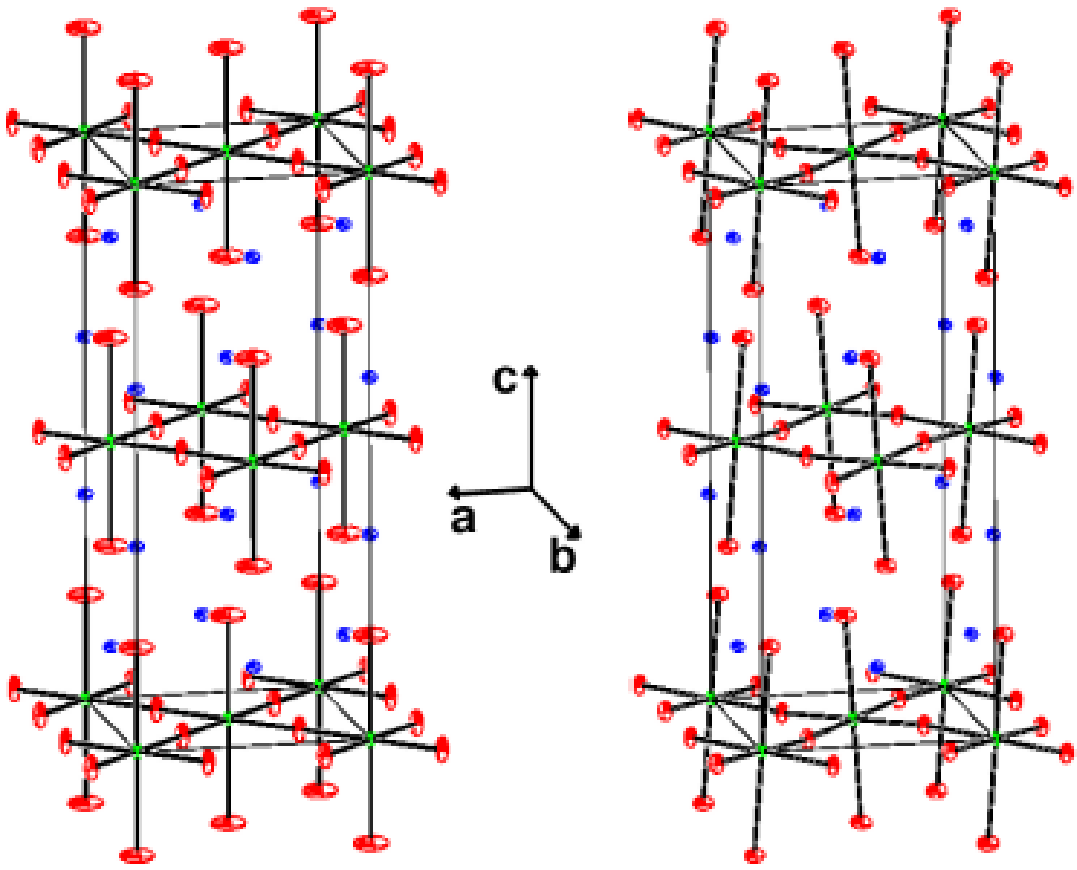


Fig. 22: *Tetragonal HTT phase (left) and orthorhombic LTO phase (right)*

In the real structure of the crystal there exist four domain types in total which are separated in pairs with Abma_1/Bmab_1 (I/II) with the (1-10) mirror plane as grain boundary and Abma_2/Bmab_2 (III/IV) with the (110) mirror plane as grain boundary, see. fig. 23.
The most accurate observation of this effect is possible only by single crystal diffraction if the diffraction plane and the *ab* plane of the sample are parallel. Be aware that the *d* spacing for the [220] reflections of all 4 domains is identical and no reflection splitting observable with powder diffraction. For the [400] reflections powder diffraction will only reveal the different *d* spacing of (400) and (040).

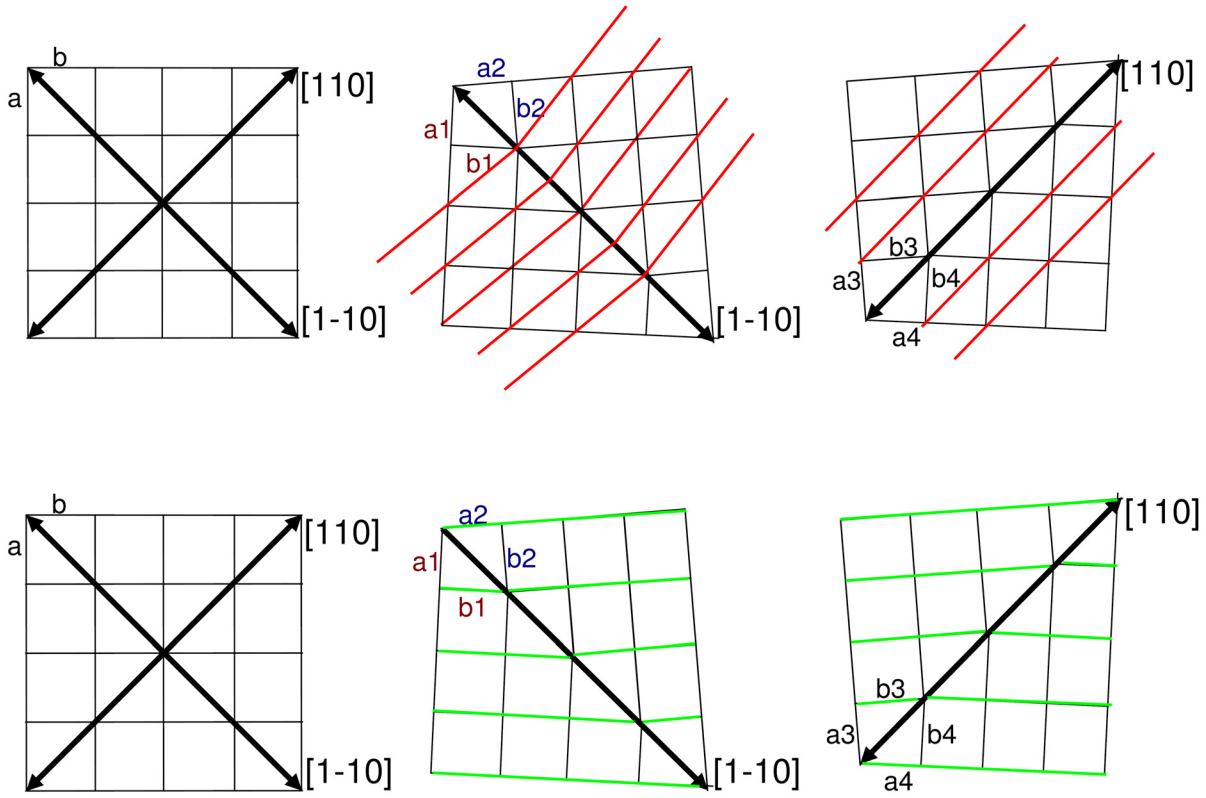


Fig. 23: *HT phase (left) and domain pair (right) with (220) planes (red) and (400) planes (green), c perpendicular*

An equal distribution of the volumetric portion of each single domain yields a ratio of intensities of 1:2:1 for the triple splitting. The distance Δω between the single peaks of a (*hkl*) reflex gives because of $(a+b)/2 = a_{1/2}$ an information about the orthorhombic *a/b* splitting. For the triple splitting of a (*hh*0) reflex is valid: $\Delta\omega = 90° - 2\arctan(b/a)$

Due to the face (=F) centring in the HTT phase only reflections with *h, k, l* of equal parity (*g* for even, *u* for odd) are allowed - (*uuu*) and (*ggg*). They are called in the following main structure reflections. The loss of symmetry in the LTO phase generates additional reflections, called super structure reflections (*e* = even index, *u* = odd index): In the A*bma* structure (*ugg*), l≠0 and (*guu*), in the B*mab* structure (*gug*), l≠0 and (*ugu*). Forbidden remain in both the HTT and the LTO phase (*uug*), (*ggu*), (*ug*0) and (*gu*0). There is no overlap between the superstructures from the A*bma* and the B*mab* domains. Therefore, although the real crystal is twinned, one can quantify the orthorhombic distortion.

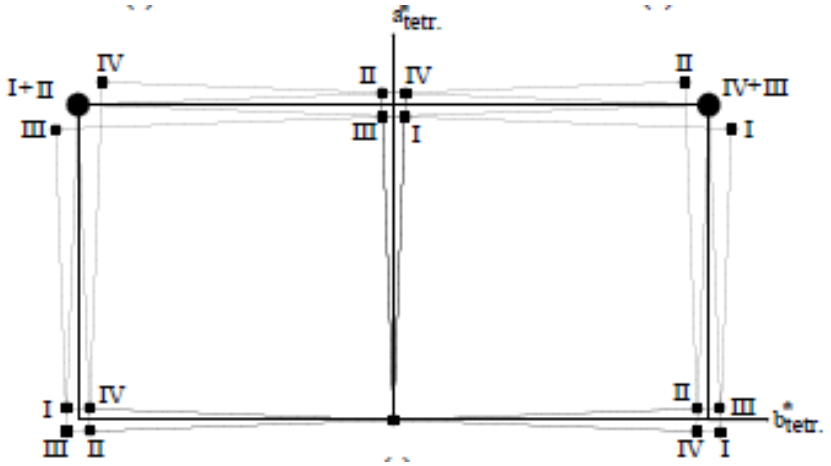


Fig. 24: *Reciprocal space (c* perpendicular) of twinned La_2CuO_4*

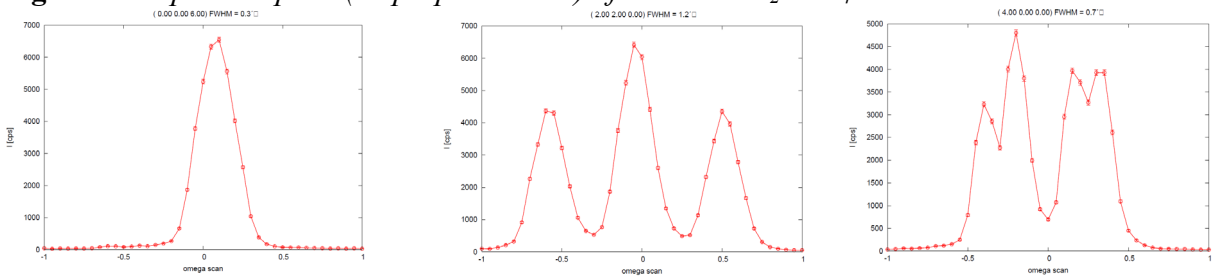


Fig. 25: *Intensity distributions of (006), (220) and (400) reflection of twinned La_2CuO_4*

The intensity contribution of the single domains corresponding to the whole intensity of a reflection can be described (taking into account the incoherent overlap of single intensities and the volumetric portions V_{A1} to V_{B2} of the domains) as follows:

$$I_{obs}(hkl) = I_{Abma1}(hkl) + I_{Bmab1}(hkl) + I_{Abma2}(hkl) + I_{Bmab2}(khl) \text{ or}$$

$$V_{total}|F_{obs}(hkl)|^2 = V_{A1}|F_{Abma1}(hkl)|^2 + V_{B1}|F_{Bmab1}(hkl)|^2 + V_{A2}|F_{Abma2}(hkl)|^2 + V_{B2}|F_{Bmab2}(hkl)|^2$$

$$= (V_{A1} + V_{A2})|F_{Abma1}(hkl)|^2 + (V_{B1}|+ V_{B2})|F_{Bmab1}(hkl)|^2$$

$$= V_{total}\ \{\alpha|F_{Abma}(hkl)|^2 + (1-\alpha)\ |F_{Abma}(khl)|^2\}$$

with α being the relative portion of the volume of A*bma* domains to the crystal..

Because of the extinction rules in the LTO phase for the super structure reflections is valid: $I_{obs}(hkl) \sim \alpha|F_{Abma}(hkl)|^2$ for A*bma* and $I_{obs}(hkl) \sim (1-\alpha|F_{Abma}(khl)|^2$ for B*mab*. Thus, one can classify directly intensities to the volumetric portions of the domain types A*bma* and B*mab* respectively. Therefore, by using one single additional parameter α to describe the relation between the twins in the structure one can determine the orthorhombic single crystal structure! This holds true although the Bragg reflections contain contributions of up to four different domains.

Light elements: The phase transition is driven by a displacement of the oxygen atoms (see fig. 22). As the oxygen atoms are much lighter than any other element in this compound, the accurate observation of these displacements depends strongly on the chosen radiation. As the atomic positions of Cu and La do not change significantly between the HTT and LTO phase the structure factor for the superstructure reflections can be written as

$$F(hkl) \sim \sum_i s_i \exp(-2\pi\imath(hx_i+ky_i+lz_i) = F(hkl)_{\text{apex oxygen}} + F(hkl)_{\text{in plane oxygen}} + F(hkl)_{\text{Remains}}$$
$$\rightarrow F(hkl)_{\text{apex oxygen}} + F(hkl)_{\text{in plane oxygen}}$$

As the apex oxygen moves away from the z=0 position to (x0z) the corresponding superstructure reflection for h uneven is

$$F(hkl)_{\text{apex oxygen}} = \sin(2\pi hx)\cos(2\pi lz) \text{ for } h \text{ odd}$$

In the case of x-rays the weak form factor of the oxygen (Z=8) against the Cu (Z=29) and La (Z=57) atoms make this intensity contribution almost invisible (<< 1% of main reflections). In the case of neutrons the scattering lengths of all atoms are in the same order of magnitude (b_O=5.803 fm, b_{Cu}=7.718 fm, b_{La}=8.24 fm) and therefore also the superstructure reflections yield easily measureable intensities significantly larger than 1% of the strongest main structure reflections.

Mean square displacements: Pure La_2CuO_4 shows a purely linear behaviour of the mean square displacements with temperature. Deviations from this harmonic behaviour of the Debye-Waller-factors can be a hint for a disturbance like an order-disorder phase transition. As the $La_{1.85}Sr_{0.15}CuO_4$ compound shows the highest T_c = 38 K it was discussed whether an order/disorder phase transition could be related to superconductivity. Bragg data sets taken with neutron single crystal diffraction at three temperatures above and below the structural phase transition (T=186 K) and the superconducting state (T=38 K) show no anomaly for all atoms including the two oxygens O1 (in-plane) and O2 (apex) (fig. 26). The only anomaly found there is the increase of U_{33}(O1) and U_{11}(O2) for *all* temperatures compared to the undoped La_2CuO_4. Harmonic lattice dynamical calculations from experimentally determined phonon dispersion curves taking into account the Sr doping were in good agreement with this observation. Thus, the random distribution of Sr atoms on La sites introduces static disorder into the structure [14].

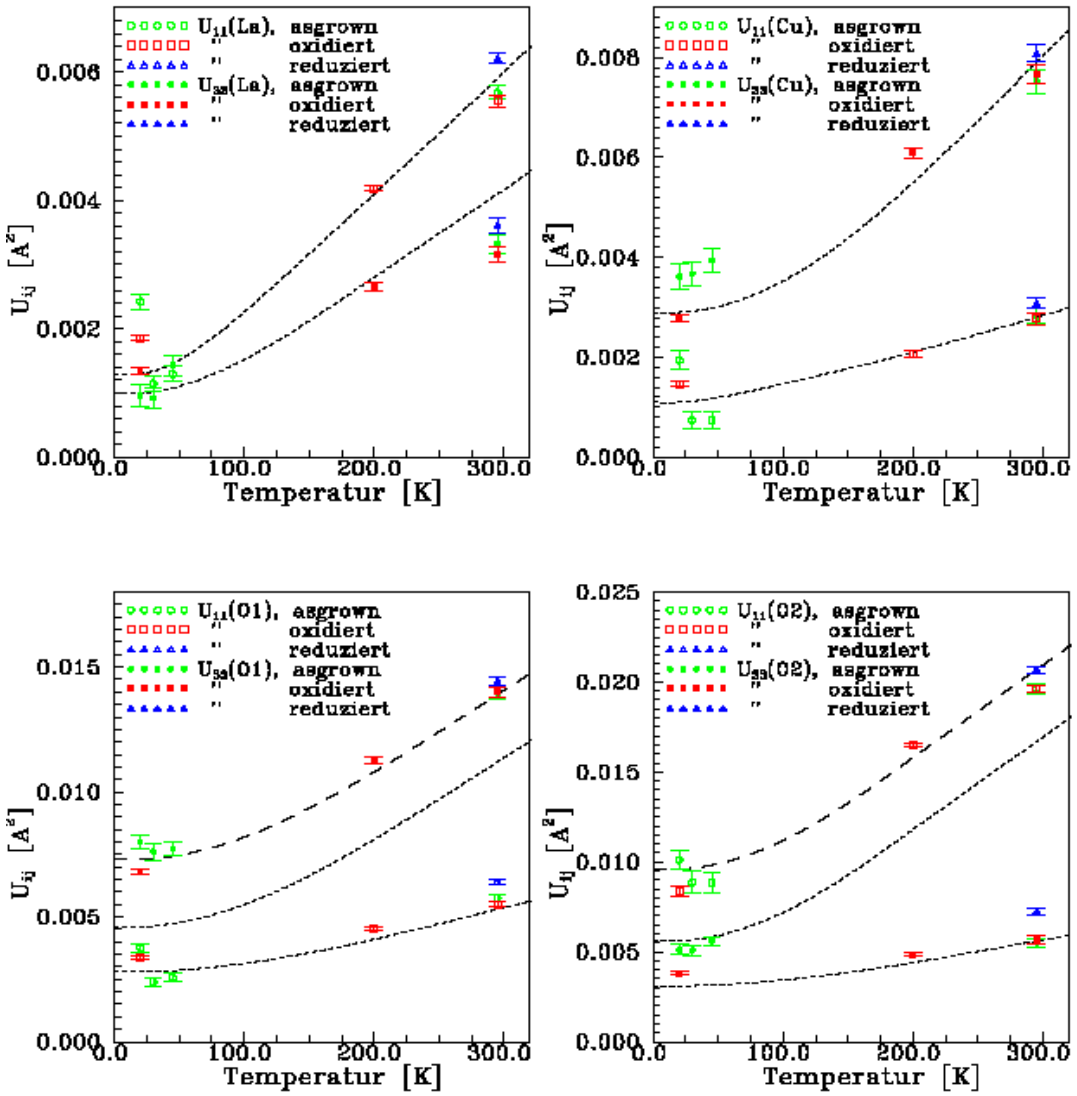
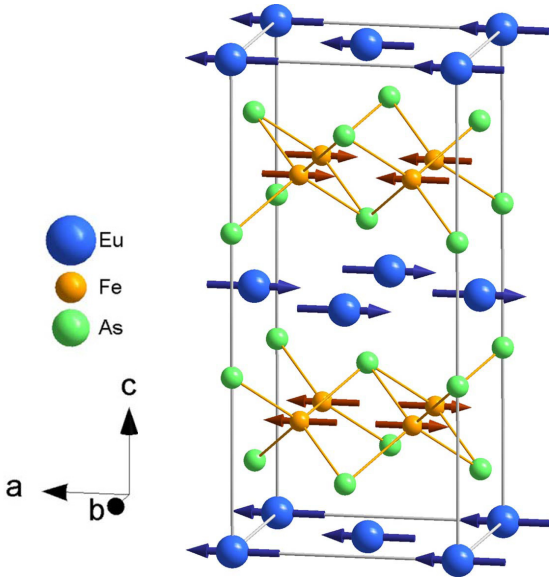


Fig. 26: *MSD of La, Cu, O1 and O2 for $La_{1.85}Sr_{0.15}CuO_4$ [14].*
The dotted lines in the middle of the diagrams for O1 and O2 are U_{33}(O1) and U_{11}(O2) of La_2CuO_4

Magnetic compounds: Some years ago FeAs compounds were found to show – like the cuprates - superconductivity if doped. It was also found that the structural properties are similar to the cuprates, e.g. a layered structure and a phase transition from a tetragonal high temperature phase (I4/mmm) to an orthorhomic low temperature phase (Fmmm, fig. 27).



An interesting member of these compounds is $EuFe_2As_2$, where doping with potassium generates superconductivity with $T_c(Eu_{0.5}K_{0.5}Fe_2As_2)$ = 31 K. The undoped compound shows antiferromagnetic (=AF) ordering of the Fe^{2+} and Eu^{2+} atoms but at different temperatures, 190 K and 19 K respectively. Neutron single crystal diffraction was used to study the nuclear *and* magnetic structures in detail as the magnetic moment of the neutron is sensitive to magnetic order (fig. 28) [15].

Fig. 27: *F/mmm phase of $EuFe_2As_2$ [15].*

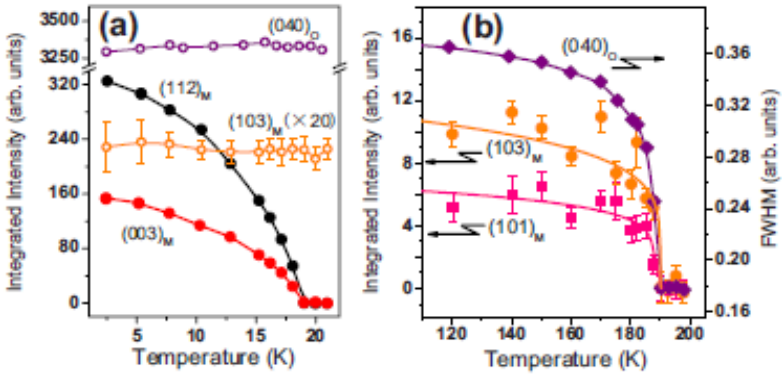


Fig. 28: *T dependent measurements of magnetic Bragg reflections (Eu^{2+} left, Fe^{2+} right)[15].*

Like for the example of La_2CuO_4 the orthorhombic structure is twinned. Careful profile analysis and Bragg data collections were used to reveal the details of the orthorhombic structure [15]. As can be seen in fig. 29 the nuclear and magnetic reflections are well separated. Additionally, the comparison of nuclear and magnetic reflections between measurement (F_{obs}) and model (F_{calc}) show good agreement for both (fig. 30).

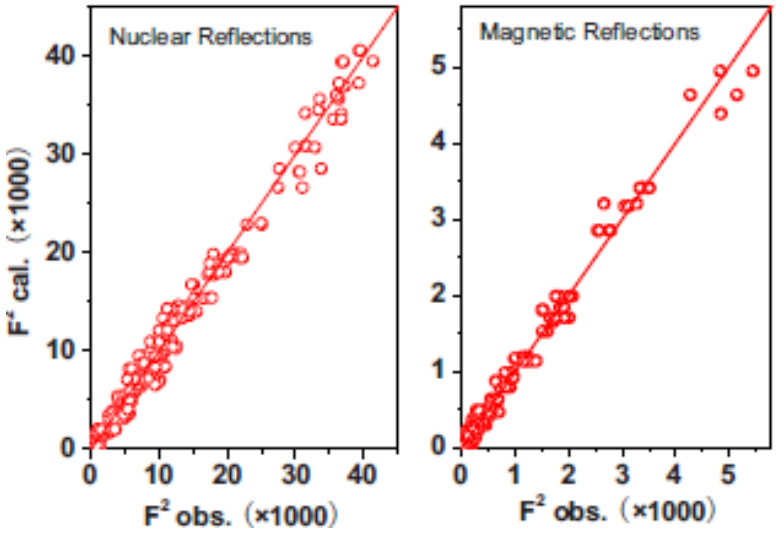


Fig. 29: *Correlation diagrams of nuclear and magnetic reflections [15].*

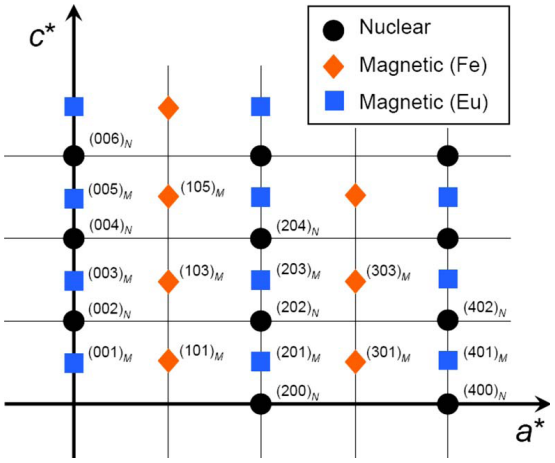


Fig. 30: *Distribution of magnetic and nuclear peaks in reciprocal space along b* [15].*

4 Summary

Powder and single crystal diffractometry are two of the most versatile tools for detailed studies on chemical and magnetic structures. The different interactions of x-ray and neutron radiation with matter allows these techniques to contribute important informations to almost every scientific area of solid state physics, chemistry, biology and material sciences. This script can offer only a short overview of the methods and their applications. Additional information can be found in the attached list of literature.
The following table presents a compact guideline which radiation and technique might be best suited to answer a specific scientific question:

	Powder	Single Crystal
X-Rays	sample mass µg	sample size µm
	very fast, in situ experiments	fast (1/2 - 2 days)
	(intrinsic & resolution dep.) reflection	very high angular resolution, anisotropic
	intensity damping by temperature factor *and* structure factor!	
	weak interaction with light elements (H, N, O, etc.)	
	no discrimination of neighbouring elements (e.g. Co, Fe, Cu)	
	large absorption effects, polarization effects	
	limited sample environment (T, H)	
	e- density maps - characterization of chemical bonds	
Neutrons	sample mass mg	sample size mm
	fast, in situ experiments	medium (3 days -7 days)
	reflection overlapp	good angular resolution
	intensity damping *only* by temperature factor	
	strong interaction with light elements, isotope specific!	
	discrimination of neighbouring elements and isotopes (H/D)	
	weak absorption effects, no polarization effects (unless pol. neutrons)	
	core and spin density maps, magnetic ordering!	

References

[1] Th. Hahn (ed.), Space-group symmetry, International Tables for Crystallography Vol. A, Kluver Academic Publishers (1995).
[2] W.H. Zachariasen, Acta Cryst. 18 703 (1965).
[3] W.H. Zachariasen, Acta Cryst. 18 705 (1965).
[4] P. Coppens and W.C. Hamilton, Acta Cryst. A 26 71-83 (1970).
[5] P.J. Becker and P. Coppens, Acta Cryst. A 30 129-147 (1974).
[6] P.J. Becker and P. Coppens, Acta Cryst. A 30 148-153 (1974).
[7] U.H. Zucker, E. Perrenthaler, W.F. Kuhs, R. Bachmann and H. Schulz J. of Appl. Crystallogr., 16 358 (1983).
[8] P. Coppens, W.C. Hamilton, S. Wilkins, M.S. Lehmann and Savariault, Datap, http://www.ill.fr/data treat/diftreat.html#single (1999).
[9] X. Turrillas, P. Barnes, D. Gascoigne, J.Z. Turner, S.L. Jones, C.J. Norman, C.F. Pygall and A.J. Dent; Radiat. Phys. Chem. Vol. 45, No. 3, pp. 491-508 (1995)
[10] M. Ermrich, F. Hahn and E.R. Wölfel, Textures and Microstructures 29, 89-101 (1997)
[11] R. Niewa, A. Czulucki, M. Schmidt, G. Auffermann, T. Cichorek, M. Meven, B. Pedersen, F. Steglich, R. Kniep, *J. Solid State Chem.* 183, 1309 (2010).
[12] J. Bednorz and K. Müller, Z. Phys. B 64 189 (1986)
[13] R.J. Birgeneau and G. Shirane, Physical Properties of High Temperature Superconductors I, Editor D.M. insberg, World Scientific (1989).
[14] M. Braden, M. Meven, W. Reichardt, L. Pintschovius, M.T. Fernandez-Diaz, G. Heger, F. Nakamura, and T. Fujita; Phys. Rev. B 63, 140510 (2001).
[15] Y. Xiao, Y. Su, M. Meven, R. Mittal, C. M. N. Kumar, T. Chatterji, S. Price, J. Persson, N. Kumar, S. K. Dhar, A. Thamizhavel, and Th. Brueckel, Phys. Rev. B 80, 174424 (2009).

Literature

N.W. Ashcroft and N.D. Mermin, Festkörperphysik, Oldenbourg 2001.
H. Ibach and H. Lüth, Festkörperphysik, Einführung in die Grundlagen, 6. Ed. Springer 2002.
C. Kittel, Einführung in die Festkörperphysik, 10. Edition, Oldenbourg 1993.
W. Borchardt-Ott, Kristallographie. Eine Einführung für Naturwissenschaftler, 6. Auflage Springer 2002.
W. Kleber, Einführung in die Kristallographie, Oldenbourg 1998
H. Dachs, Neutron Diffraction, Springer (1978)
D.J. Dyson, X-Ray and Electron Diffraction Studies in Material Science, Maney Pub 2004.
C. Giacovazzo, Fundamentals of Crystallography, 2nd Ed., Oxford University Press 2002.
L.A. Aslanov, Crystallographic Instrumentation, Oxford University Press 1998.
M.T. Dove, Structure and Dynamics. An Atomic View of Materials, Oxford University Press 2003.
W. Clegg, Crystal Structure Analysis. Principles and Practice, Oxford University Press 2001.

D 4 Inelastic scattering: Lattice, magnetic and electronic excitations.

Jörg Voigt

Jülich Centre for Neutron Science

Forschungszentrum Jülich GmbH

Contents

[0]Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 What do we want to measure?

In a former lecture, lattice vibrations (phonons) and magnetic excitations (magnons) in solids have been derived theoretically. In the present lecture I will show, how such collective excitations can be measured.

The thermal properties of a material depend on how the constituents move on a microscopical scale. How the atoms/molecules move determines e.g. how heat is transported, what happens at phase transition and how lattice, spin, electronic etc. degrees of freedom interact... The present lecture emphasizes the measurements of lattice and spin excitations. Electronic excitations are only covered as they are involved in the resonant inelastic x-ray scattering process, which is an emergent technique to study magnetic excitations with x-ray.

Phonons and magnons are characterized by their dispersion relations, i.e. $\hbar\omega = f(\mathrm{q})$. Here $\hbar\omega$ denotes the Eigenvalue of the Hamiltonian and $\hbar$q the momentum of the excitation. Hence the probes we use for spectroscopy must exchange energy and momentum with the system under study. Typical excitation energies range from a few meV up to several 100 meV. The momentum is governed by the interatomic distances in the sample, which are on the nm scale. Probes that carry the respective momentum are e.g. photons with an energy of several keV, neutrons in the thermal energy range, electrons or α particles. So if the probes are scattered by the sample, the momentum transfer tells us, where the atoms in the material reside. The energy transfer tells us about the bonding of an atom. If the experiment resolves both momentum and energy transfer, we get the detailed information about how the atoms can move in a crystal, i.e. the information about the dispersion relation, which can then be compared to theoretical models describing the sample under study. Light scattering methods, e.g Raman or Brillouin scattering, are employed to investigate elementary excitations in the Brillouin zone centre. Electrons and ions can be used as probes for near surface properties, which may coincide with the behavior in the bulk of a material or may exhibit additional interesting features due to the loss of translational symmetry. This lecture covers inelastic scattering of neutrons and x-ray photons as bulk probes, which penetrate deeply into the sample and provide access to the dispersion relation across the whole Brillouin zone.

1.1 A 'simple' dispersion

Fig. 1 shows a 'simple' dispersion relation, as it has been derived in the lecture of Karin Schmalzl (B4). The dispersion relation for lattice excitations describes, how the displacements of the atoms from their equilibrium position change the energy of the crystal. As the atoms are bonded inside the crystal, the displacement of a single atom forces the other atoms to be displaced, too. The symmetry and the bonding details will determine how the displacement pattern will then look like. In the quantized picture we call such an excitation of the whole lattice a phonon.
Due to the translational symmetry, the phonon can be labeled by its momentum, which is proportional to the reciprocal lattice vector that describes the displacement pattern. All possible phonons can be expressed by reciprocal lattice vectors in the first Brillouin zone grace to the discrete nature and the translational symmetry of the real lattice. From the experimental point of view one may therefore choose a certain reciprocal lattice point and study the Brillouin zone starting from this centre called Γ point.

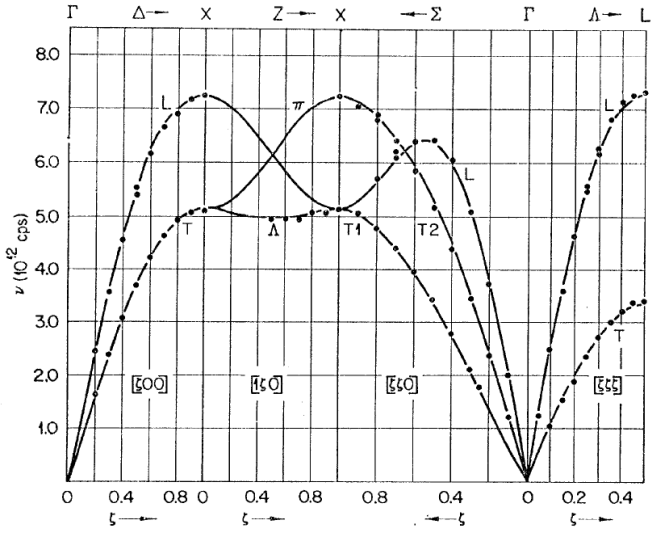


Fig. 1: *a) Dispersion relation for bcc copper along different main symmetry directions. b) Scattering plane in reciprocal plane. Indexed dots refer to Bragg points of the fcc Cu. Two scattering triangles spanned by the initial and final wave vector k and k′ are indicated, one for the measurement of transverse modes and one for longitudinal modes. The dashed triangle in the upper right indicates the nomenclature for high symmetry points in the Brillouin zone. Taken from [1]*

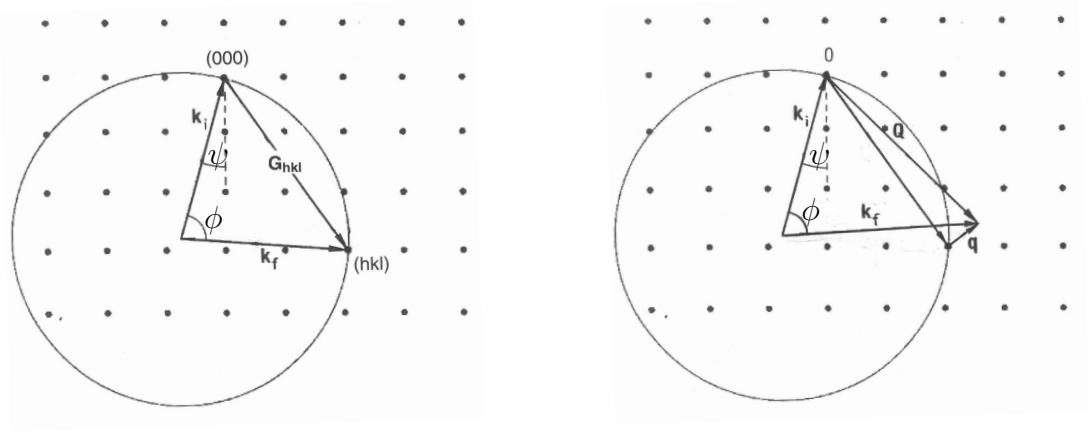
.

Similar dispersion relations exist for the quantized excitations of the spin system. The that case one can only expect a response of the sample, if the probe interacts with the magnetic moments. For neutrons this interaction is comparable to the nuclear interaction, while for x-ray it is only a relativistic correction to the dominant charge scattering. However, resonance effects can acquire sensitivity to magnetic moments via spin-orbit coupling as will be discussed in sec. 4.3. To find out, where the atoms are and how they move in a sample, the probe excites phonons or magnons, respectively. The scattering obeys energy, momentum and angular momentum conservation. Hence we need to measure the momentum, energy and possibly angular momentum transfer onto the probes in an actual scattering experiment. In sec. 2 the relations between probe and the sample momentum and energy will be derived, before I explain in sec. 3, how the measurement is actually done. The details of the interaction between sample and probe will influence the scattered intensity and further characterize the properties of the excitation as detailed in sec. 4.

2 Scattering

2.1 Kinematics

Elastic scattering It is instructive to start the explanation of energy and momentum transfer with the visualisation of a diffraction experiment (elastic scattering) in reciprocal space. Let us take a look at the reciprocal space of a simple cubic lattice. In Fig. 2 a part of the plane perpendicular to a [100] direction is shown, which exhibits a quadratic arrangement of reciprocal lattice dots. They represent possible Bragg reflections. If the crystal is perfect except for a



a) b)

Fig. 2: *a) Reciprocal space and vector representation for elastic scattering: the $\{100\}$ zone of a simple cubic lattice showing Ewald's construction for Bragg reflection. The number (letters) in brackets denote the Miller indices of the reciprocal lattice vector $\boldsymbol{\tau}$. b) Reciprocal space and vector representation for inelastic scattering: The momentum transferred to the sample is the sum of the reciprocal lattice vector $\boldsymbol{\tau}$ and the momentum transfer q to a phonon within the Brillouin zone around $\boldsymbol{\tau}$.*

certain degree of mosaicity (micro crystallites are all aligned within a small deviation η from a given orientation), the elastic scattering is confined to the Bragg dots. For elastic scattering is $|\mathrm{k}'| = |\mathrm{k}| = \mathrm{k}$. The circle with radius k (the Ewald circle, respectively sphere in three dimensions) contains the origin of the reciprocal space and at least one other point of the reciprocal lattice of the crystal. The circle contains all scattering vectors $\mathrm{Q} = \mathrm{k}' - \mathrm{k}$ for a fixed k. The condition

$$\mathrm{Q} = \boldsymbol{\tau} = \mathrm{k}' - \mathrm{k} \tag{1}$$

is nothing more than the familiar Bragg condition expressed in vector notation. The norm of eq. 1 gives

$$\begin{aligned} |\mathrm{Q}| = |\boldsymbol{\tau}| = \frac{2\pi}{d_{hkl}} = 2k\sin\theta = \frac{4\pi}{\lambda}\sin\theta \\ \Leftrightarrow n\lambda = 2d_{hkl}sin\theta. \end{aligned} \tag{2}$$

In this equation d_{hkl} denotes the interplanar spacing of the lattice planes with Miller indices h, k, l. θ is the glancing angle of reflection from these planes. n can take any integer value. Hence integer fractions of the fundamental wavelength $(n = 1)$ satisfy the reflection condition as well. By adjusting the the scattering angle $2\theta = \phi$ between k and k' the norm of Q is controlled. The orientation of Q is set by the sample rotation ψ. Thus, any point in reciprocal space can be reached by an appropriate choice of the incident wave vector k, the scattering angle ϕ and the sample orientation ψ relative to k.

Inelastic scattering For inelastic scattering, the situation is more complicated. Since energy is transferred between the sample and the neutron, the norm of k and k' is different. In that case,

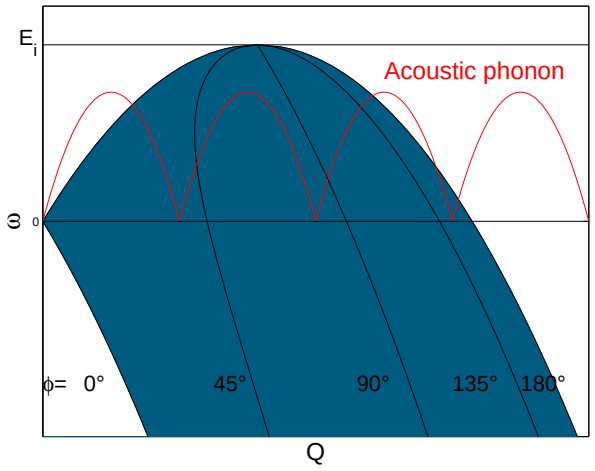


Fig. 3: *Relation between norm of the momentum transfer $\hbar Q$ and $\hbar\omega$ for different values of the scattering angle ϕ for neutron scattering. The blue area encloses the accessible Q, ω space. Black lines indicate the scattering angle Φ. As detectors cannot cover all angles, the region measured is slightly smaller. Positive values of ω correspond to energy loss of the neutron. Acoustic phonon branches in several Brillouin zones are shown as well. Note: For small momentum transfer, a phonon cannot be observed, if the sound velocity within the sample is higher than the neutron velocity (the velocities are given by the slope of the respective curves at $Q = 0$). For x-ray scattering the whole area is accessible, as the photon energy is 4 to 5 orders of magnitude larger than the excitation energy.*

the diagram 2a) has to be modified as shown in Fig. 2b). Writing down momentum and energy conservation gives:

$$\begin{aligned} \mathrm{Q} &= \mathrm{k}' - \mathrm{k} && (3)\\ \mathrm{Q}^2 &= \mathrm{k}^2 + \mathrm{k}'^2 - 2\mathrm{k}\mathrm{k}'\cos\phi && (4)\\ \hbar\omega &= E_i - E_f = \frac{\hbar^2}{2m}(\mathrm{k}^2 - \mathrm{k}'^2)(\text{for neutrons}) && (5)\\ \hbar\omega &= E_i - E_f = \hbar c(\mathrm{k} - \mathrm{k}')(\text{for x-ray}) && (6) \end{aligned}$$

Combining these equations yields the accessible region in the Q, ω space for a given incident energy E_i, which is limited for neutron scattering (see Fig. 3). The highest accessible energy transfer for a given momentum transfer is given by the cases of forward scattering ($\phi = 0$) and exact backscattering ($\phi = 180$), which can not exactly be accessed in an actual instrument.
The transferred energy $\hbar\omega$ depends in a periodic lattice only on the relative momentum $\hbar$q defined within a Brillouin zone. Therefore the momentum transfer is referred to the nearest reciprocal lattice vector.

$$\mathrm{Q} = \boldsymbol{\tau} + \mathrm{q} \qquad (7)$$

as depicted in Fig. 2b). An acoustic phonon branch has been included in Fig. 3 for several Brillouin zones. Note that an excitation at the origin of the reciprocal space can only be measured, if the sound velocity within the material is smaller than the neutron velocity. This limitation is lifted for the case of inelastic x-ray scattering.

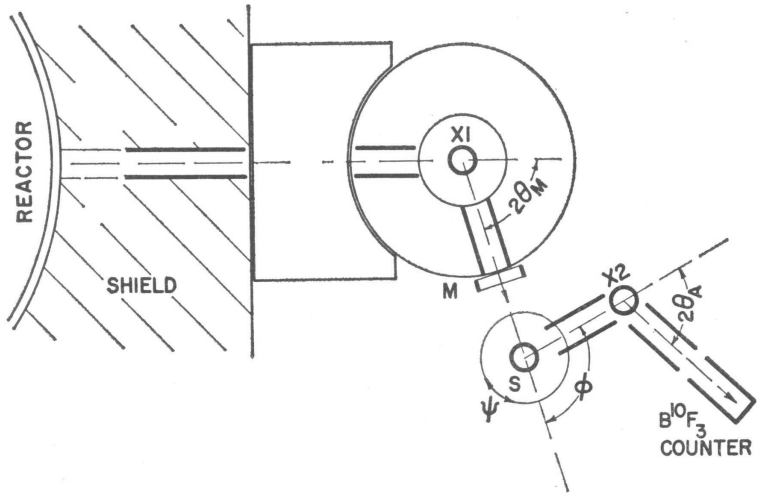


Fig. 4: *Schematic of a triple axis spectrometer. The angle Φ denotes the scattering angle between k, k', the angle ψ gives the orientation of the scattering triangle with respect to the reciprocal lattice of the sample (cp. Fig 1 b). The initial and final wavelength is defined by the scattering angles $2\theta_{M,A}$, respectively.*

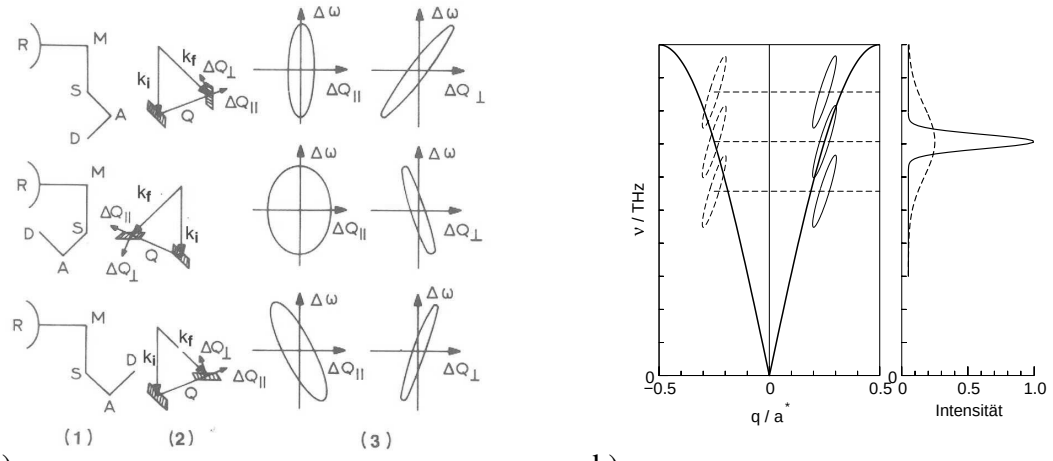
3 Methods

3.1 The neutron three axis spectrometer

The most common instrument using Bragg diffraction to prepare and analyse the neutron wavelength is the three axis spectrometer (TAS). In 1994 the Nobel price has been awarded to its inventor Bertram Brockhouse "for pioneering contributions to the development of neutron scattering techniques for studies of condensed matter" and in particular "for the development of neutron spectroscopy".
The particular instrument name stems from the existence of three axes around which important components of the instrument can be rotated (see Fig. 4). In order to select the desired wavelength according to eq. 2, the monochromator as well as the sample table have to be rotated around the first axis in a $\theta - 2\theta$ mode (X1 in Fig. 4). Around the second axis (the sample axis) the analyser/detector unit can be rotated, whereby the sample scattering angle ϕ is selected with respect to the direction of the incoming neutron wave vector k. Around the third axis (X2 in Fig. 4), both the analyser crystal and the detector are rotated in a $\theta - 2\theta$ mode in order to probe the wavelength (energy) of the scattered neutrons. An additional rotation around the sample axis (angle ψ) allows to scan any desired path within the reciprocal space plane spanned by k and k$'$.

The resolution of a triple axis instrument depends on the relative sense of rotation at the instrument spectrometer axes, as shown in Fig. 5a). The resolution function of a TAS has been derived analytically by Cooper and Nathans [2] and by Popovici [3]. The resolution of the final and initial wave vector depends on the properties of the monochromator and analyzer crystals, respectively:

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta d_{hkl}}{d_{hkl}} + \cot\theta\Delta\theta \tag{8}$$



a) b)

Fig. 5: *a)Resolution volumes of a TAS in different geometries. The hashed areas indicate the deviation of k, k' from their most probable valuesFrom [1]. b) Effect of focusing: If the orientation of the resolution ellipsoid is adopted to the slope of the dispersion, the peak shape of the inelastic signal can be optimized.*

with d_{hkl} denoting the lattice spacing of the (hkl) reflection and θ defined by the scattering angle 2θ. The first contribution to the resolution is mainly determined by the primary extinction length, which describes, how deep the neutron beam penetrates the crystal, when it is in reflection position. Because every mirror plane reflects a part of the incoming intensity, this length is limited to several 10^4 lattice spacings, giving a resolution for a perfect single crystal, which is about 10^{-4}. The second contribution originates from the mosaic spread in a real crystal. It grows for smaller scattering angles 2θ and vanishes, if 2θ approaches 180°. Since the neutron guides, which transport the neutrons from the reactor core to the instrument, determine the divergence that arrives at the monochromator within 1°, a very high resolution would yield only vanishing intensity passing the monochromator. Therefore typical monochromator crystals for neutrons have a rather wide mosaic spread or an artificial variation of the lattice constant. Since the energy of a thermal or cold neutron is on the order of several meV, typical excitations in a solid can be resolved by a 10% resolution of the initial and final energy. The situation is drastically different, if we talk about inelastic x-ray scattering. Typical photon energies are on the order of several keV. Accordingly a resolution power of 10^{-7} is required. Since x-ray beam from a third generation synchrotron is extremely bright, one can effort to minimize every contribution to the resolution and realize an intense beam, that allows the measurement of an inelastic signal.
By an appropriate choice of the scattering geometry, the resolution ellipsoid can be oriented nearly parallel to the dispersion. In that case the folding of the scattering function with the resolution function will be as narrow as possible (cp. Fig 5b)).

3.2 Multiplexing TAS

As explained in the last section three axis spectrometers can be tuned with optimized resolution to any position inside the scattering parabola. To survey the reciprocal space it is then necessary to scan along lines. This can be very time consuming, especially when the signals are weak and every point must be measured for a long time. To improve the mapping capabilities different

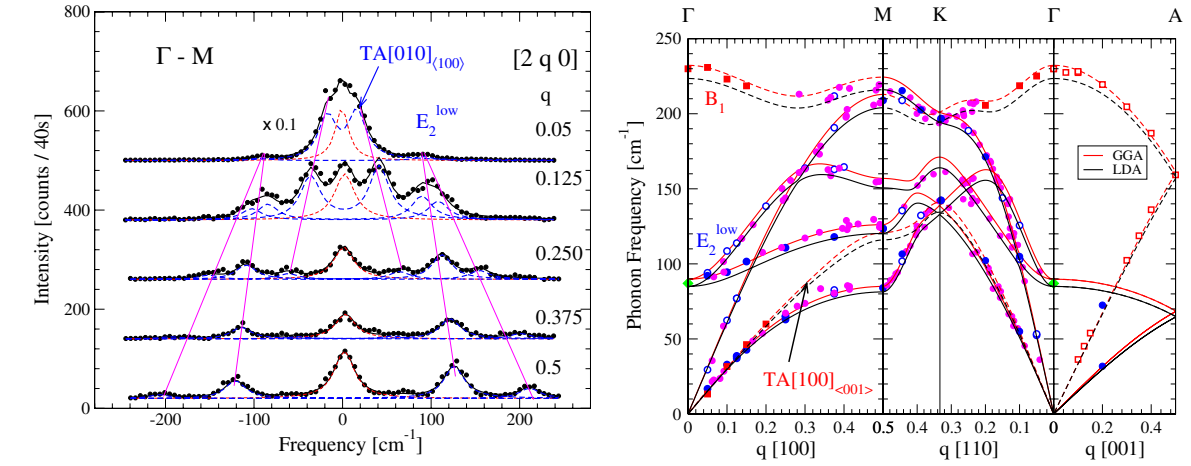
multiplexing schemes have been developed [4, 5, 6, 7]. They all have in common that not a single k′ is analyzed but that many analyzer crystals are arranged around the sample to record not a point but a line in reciprocal space. The multiplexing setups are addons to a conventional TAS. They are not as flexible, but they provide the possibility to cover large areas in reciprocal space. Special regions of interest may then be focused by conventional TAS with optimized setting of the resolution conditions with respect to the scientific question.

3.3 Inelastic x-ray scattering

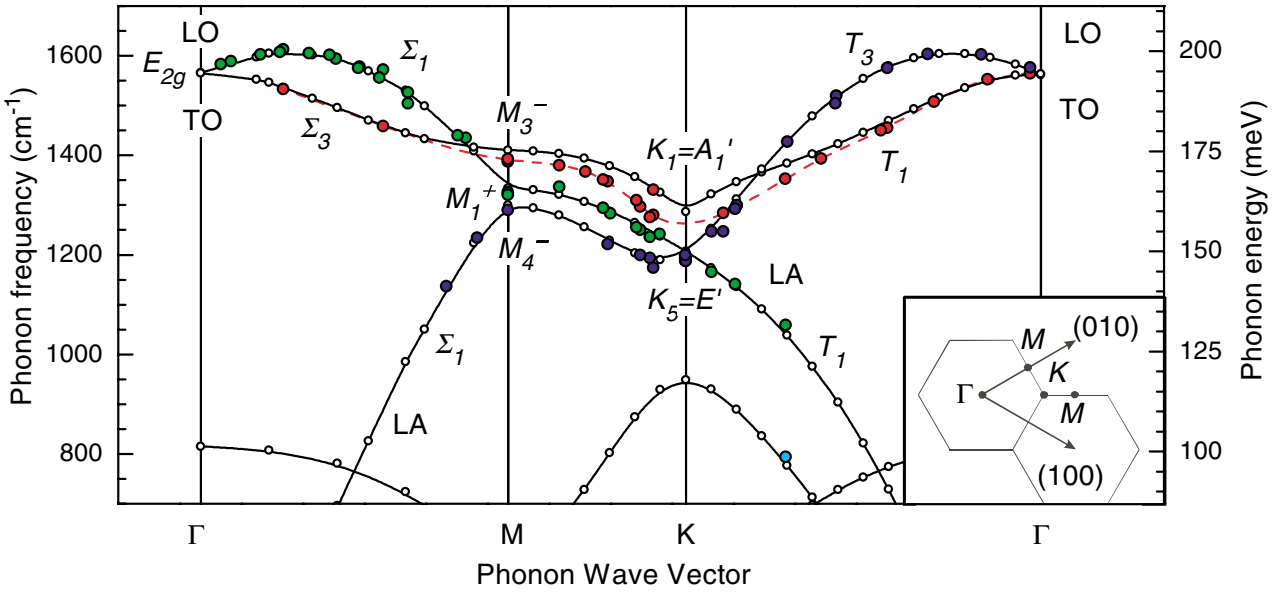
While inelastic neutron scattering is the traditional method to study coherent excitations in matter, there exist several applications that are limited by the neutrons flux of todays neutron sources. Fig. 6a shows an example of IXS from thin films. The high brightness of the synchrotron source provides enough photons to be scattered by the phonons of the thin films. Furthermore the higher absorption of the photons prevents the scattering from the substrate, which causes a strong inelastic signal in the case of the deeper penetrating neutrons. Fig. 6b is an example of high energy phonons, which can easily be accessed by x-ray scattering. If one wants to study such high energies with neutrons either a hot moderator (e.g. Graphite at 2000 K) or a short pulse spallation source with a large number of undermoderated neutrons is required.

Measurements of lattice and excitations with energies $\hbar\omega = 1 - 1000$ meV with photons of energy $E_i = 10 - 50$ keV requires a relative energy resolution better than 10^{-6}. This can be achieved using backscattering from perfect crystals. The setup of the inelastic x-ray scattering beamline ID28 at the ESRF is shown in Fig. 7.
After premonochromatization, the initial energy is defined by a flat backscattering monochromator. The penetration of a x-ray beam in a non-absorbing perfect crystal generally referred to as primary extinction length depends on the reflectivity of the single lattice plane of the considered reflection. The higher is the reflectivity, the smaller is the penetration. Going to reflections with large Miller indices increases the resolving power: this is a direct consequence of the reduced form factor for large |Q|. A $(13\ 13\ 13)$ reflection from a perfect Silicon crystal reaches $\Delta E/E = 10^{-8}$. This resolution has to be matched by the second term in eq. (8) $\cot\theta\Delta\theta$. Taking into account the angular divergence delivered from a third generation synchrotron source ($\Delta\theta \approx 10\mu$Rad) $\cot\theta$ should be smaller than 10^{-4} ($\theta \approx 89.98°$).
After further optical components, the photons are scattered from the sample. The energy resolution of the analyzer should be the same as the monochromator. However, the angular acceptance is defined by the desired Q resolution, which is relaxed as compared to the energy resolution. Typical values for ΔQ range from 0.2 to 0.5 nm^{-1}. This corresponds to an angular acceptance for the analyzer of about 10 mrad. The angular acceptance is obtained by the use of a focusing system. To preserve crystal perfection properties small flat perfect crystals are layed on spherical surface. The curvature is defined by a 1:1 pseudo *Rowland* geometry with aberrations kept such that the energy resolution is not degraded.
The backscattering conditions requires an alternative way to scan the energy transfer. Since the angle is fixed, one has to vary the lattice constant according to the Bragg eq. 2. The energy transfer is determined by the relative temperature between monochromator and analyzer. Considering $\Delta d/d = \alpha\Delta T$, with $\alpha = 2.56 \cdot 10^{-6}\mathrm{K}^{-1}$, the temperature has to be controlled with a precision of 0.5 mK to obtain an energy step of about 1 tenth of the energy resolution. The overall resolution of the spectrometer at ID28 is shown in Fig. 8.

To overcome the constraints set by the backscattering geometry alternative high resolution



(a) Phonon dispersion of 6.2 μm thin InN film on GaN substrate. From [8]



(b) Phonon dispersion of a graphite single crystal. From [9]

Fig. 6: *Examples of phonon studies with inelastic x-ray scattering.*

monochromators have been developed for high resolution synchrotron applications (see [10] and references therein). These monochromators increase the energy resolution by the use of asymmetric Bragg reflections to achieve a higher energy bandwidth than the intrinsic resolution of a single Bragg reflection [11] (see Fig. 9). Such monochromators allow a larger energy range to be scanned. Thermal stability is very important in order not to change the lattice constants. It is furthermore beneficial, when the lattice expansion vanishes at the operation temperature.

3.4 Time-of-flight spectroscopy

In contrast to x-ray photons the neutron is a massive particle and the velocity depends on the energy. Typical velocities for thermal and cold neutrons range from $v_n = 3956\frac{\mathrm{m}}{\mathrm{s}}$ for neutrons with $\lambda = 1$ Å to $v_n = 395.6\frac{\mathrm{m}}{\mathrm{s}}$ for $\lambda = 10$ Å. The energy is of course proportional to the square

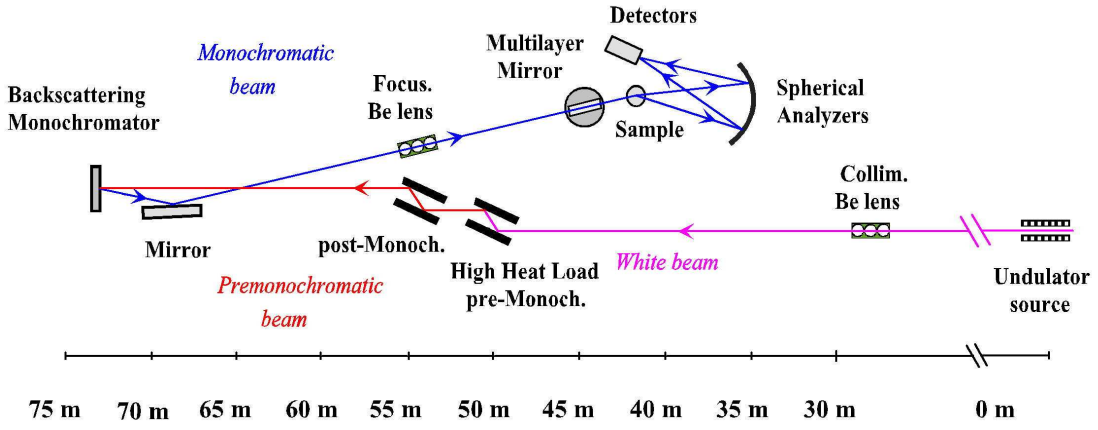


Fig. 7: *Setup of the inelastic scattering beamline ID28 at the ESRF.*

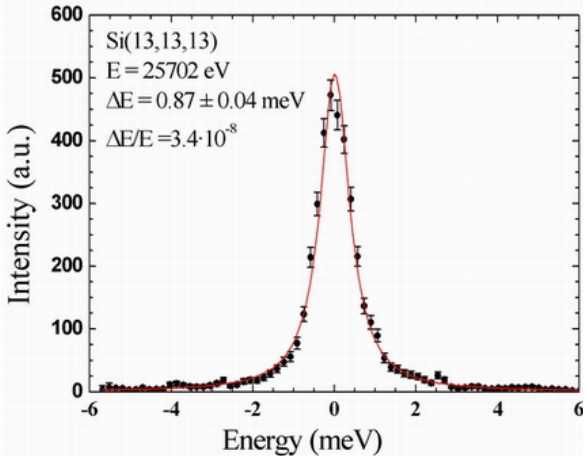


Fig. 8: *Instrumental resolution function of ID28 at the ESRF determined by scattering from a Plexiglas sample at Q=10 nm^{-1} and T=10 K.*

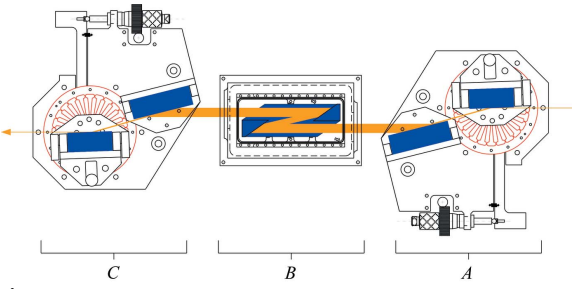


Fig. 9: *Design of a six-reflection cryogenically stabilized high-resolution monochromator. The x-ray beam enters from the right. From [10].*

	Neutrons	X-rays
λ[Å]	0.2 – 20	0.5 – 40
E_i[eV]	$0.0818 \cdot \lambda^{-2}$	$12398 \cdot \lambda^{-1}$
v[m/s]	$3956 \cdot \lambda^{-1}$	c
$\Delta\hbar\omega$[meV]	10^{-3} – 10	50 – 500
Divergence [rad]	0.01	10^{-6}
Sensitivity		
Phonons	+	+
Magnons	+	(non resonant) - (RIXS) +
Electronic excitations	-	+

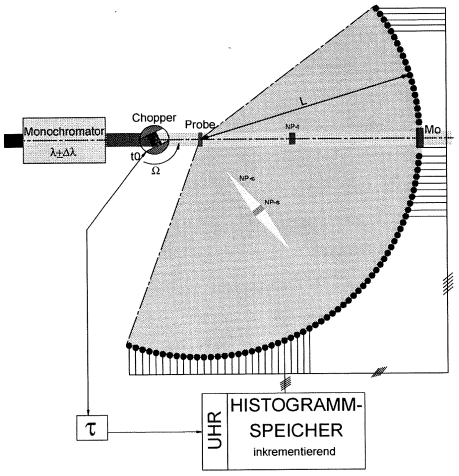
Table 1: *Comparison Neutron ↔ x-ray in the wavelength range that is used for spectroscopy.*

of the velocity and inverse of the wavelength λ:

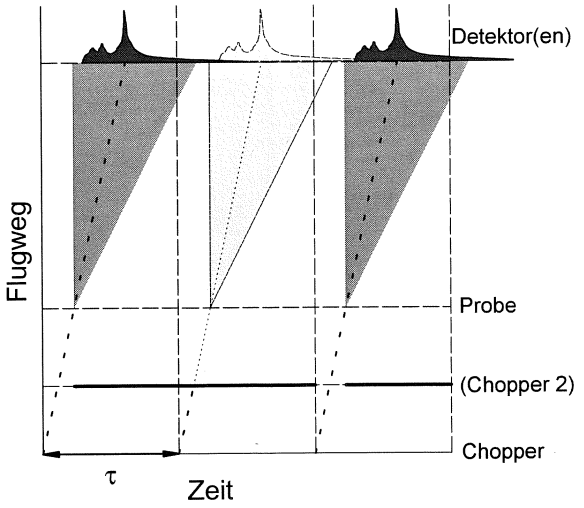
$$E_n = \frac{1}{2} m_n v_n^2 \tag{9}$$

$$= \frac{h^2}{m_n \lambda^2}. \tag{10}$$

Accordingly 1 Å neutrons have an energy of 81.8 meV and 10 Å neutrons have 0.818 meV. This offers an alternative way to measure the energy exchange with the sample: if we define the initial neutron velocity by an appropriate monochromator and define a precise start signal for the neutrons hitting the sample by a chopper, we can analyze the energy exchange to the sample by recording the arrival of the neutrons at the detector in a given distance from the sample and the chopper. In contrast to three axis spectroscopy, the energy can then be determined simultaneously for all scattering angles ϕ. Therefore by covering a large solid angle with detectors, one could measure the complete Q, ω space very effectively. However, the decoupling of energy and scattering angle comes for a price: one can use only neutrons in a short burst time $\tau < 20\mu$s and has then to wait until all neutrons arrive at the detector to record the full energy spectrum, before the next pulse can start, see Fig. 10b. So for a time-of-flight instrument at a continuous source, one can use the neutron beam during 1-2% of the time. At short pulse spallation sources, all neutrons are created in a short pulse. For time-of-flight experiments on can therefore us the full peak flux and gain over constant beam sources a factor ≈ 100, even if the time averaged flux is comparable or lower. The intensity scattered into the solid angle $d\Omega$ defined by the desired Q resolution is still sometimes at the limit of detection. Traditionally time-of-flight spectroscopy is used mainly to study excitations that are not sharply defined in momentum transfer. Then the intensity can be integrated over a larger solid angle. One example is the incoherent one phonon scattering cross section as a measure for the phonon density of states, as it integrates all scattering vectors Q for a given energy ω. Crystalline electric field excitations that depend mainly on the local environment are another typical application for time-of-flight spectroscopy. With the new Megawatt spallation sources but also with new instruments at reactor sources, as the new TOPAS instrument at the FRM2, the flux limit will be lifted and the full potential of TOF spectroscopy will be available for mapping the excitation landscape in condensed matter research.



(a) Generic setup of a time-of-flight spectrometer.



(b) Path-time diagram for a configuration with frame overlap choppers

Fig. 10: *ToF spectroscopy.*

4 Response

So far we have discussed, how we can tune an neutron or x-ray spectrometer to measure a certain energy and momentum exchange with the sample. Often this is also the major aim of an experiment. Important insight into the dynamics of a system can often be acquired from the knowledge of the dispersion relation. But a scattering experiment can provide more information, which is contained in the scattered intensity.

The inelastic scattering of neutrons and x-ray is usually described in the Born approximation. The information on the dynamics is contained in the scattering function $S(\mathrm{Q},\omega)$, which is connected to the double differential cross section $\frac{d^2\sigma}{d\Omega\omega}$, the basic quantity in a scattering experiment [12]:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}\omega} = A(\mathrm{Q})S(\mathrm{Q},\omega). \tag{11}$$

The properties of the scattered particle as well as the interaction potential with the sample are included in the factor A.
The scattering function $S(\mathrm{Q},\omega)$ is the Fourier transform of the pair correlation function, introduced in the lecture of R. Zorn (A5). It depends only on the momentum and the energy transferred to the sample and not on the actual values of k and k′. It contains the information on both the positions and motions of the atoms comprising the sample. Via the dissipation-fluctuation theorem, the scattering function is connected to the dissipative or imaginary part of the dynamical susceptibility [13]:

$$S(\mathrm{Q},\omega) = \frac{\chi''(\mathrm{Q},\omega)}{1-\exp(-\hbar\omega)/(k_B T)} \tag{12}$$

The measured intensity in an actual inelastic scattering experiment is the folding of the resolution function of the instrument with the scattering function in Q,ω space:

$$I(\mathrm{Q},\omega) \propto \iint R((\mathrm{Q}'-\mathrm{Q}),(\omega'-\omega))S(\mathrm{Q},\omega)d\mathrm{Q}'d\omega' \tag{13}$$

The resolution function R depends on the are a probability distributions $p_{i,f}$ for the initial and final wave vectors:

$$R(\mathrm{Q},\omega) = \iint p_i(\boldsymbol{k_i'}) p_f(\boldsymbol{k_f'}) \delta(\mathrm{Q'} - (\mathrm{k'} - \mathrm{k})) \delta(\omega' - \frac{\hbar}{2m}(\mathrm{k}^2 - \mathrm{k'}^2) d\boldsymbol{k_i'} d\boldsymbol{k_f'} \tag{14}$$

The normalisation of the resolution function is the product of the volume elements in reciprocal space defined by probability distributions:

$$\iint R((\mathrm{Q'} - \mathrm{Q}), (\omega' - \omega), \mathrm{Q}, \omega) d\mathrm{Q'} d\omega' = V_i(\mathrm{k}) \cdot V_f(\mathrm{k'}) \tag{15}$$

The intensity one can expect for a given Q, ω thus is proportional to this product. These distributions are defined by the collimation on the primary side (before the sample) and the secondary side (between the sample and the detector). Hence it becomes clear that intensity and resolution (defined by the collimation) have to be traded according to the present scientific problem.

4.1 Single phonon scattering

The double differential cross section for coherent single phonon scattering of neutrons contains more information about the system under study than only the dispersion relation, which is contained in the delta function eq. (19) [14]:

$$\begin{aligned}
\left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega \mathrm{d}\omega}\right)^{\pm}_{coh} &= \frac{\mathrm{k'}}{\mathrm{k}} \frac{(2\pi)^3}{v_0} \sum_{\boldsymbol{\tau}} \sum_{j,\mathrm{q}} \\
&\cdot |\sum_i \frac{b_i}{\sqrt{m_i}} \exp[-W_i(\mathrm{Q}) + i\mathrm{Q} \cdot \boldsymbol{R_i}](\mathrm{Q} \cdot \boldsymbol{e_i^j})|^2 && (16)\\
&\cdot \omega_j^{-1} && (17)\\
&\cdot (n(\omega_j(\mathrm{q})) + \frac{1}{2} \pm \frac{1}{2}) && (18)\\
&\cdot \delta(\omega \mp \omega_j(\mathrm{q})) \delta(\mathrm{Q} \mp \mathrm{q} - \boldsymbol{\tau}) && (19)
\end{aligned}$$

In the expression for the double differential x-ray scattering cross section, the nuclear scattering length must be replaced by the atomic form factor as long as the photon energy is not close to an absorption edge, but the structure of the cross sections remains the same.
Expression 16 is called the dynamical structure factor, in analogy to the structure factor in elastic scattering. Its strength depends on the nuclear coherent scattering length of the constituting nuclei b_i, which is the mean value of the different scattering length for different isotopes or different nuclear spin. It includes the Debye-Waller factor $\exp[-W], W = \langle \mathrm{Q} \cdot \boldsymbol{u} \rangle^2$, introducing the displacements $\boldsymbol{u}$ of atoms from their equilibrium position.
The intensity is proportional to a factor $|\mathrm{Q} \cdot \boldsymbol{e_i^j}|^2$, with the polarization vector of the phonon mode $\boldsymbol{e_i^j}$. This means, it becomes favorable to measure phonons with large Q. This property provides an opportunity to distinguish between magnetic and lattice excitations. As we will see later, scattering from magnetic excitation becomes less probable for large momentum transfer due to the magnetic form factor. Therefore measurements in different Brillouin zones tell you, whether your signal is of magnetic or nuclear origin.
Due to the scalar product in eq. (16) only if the displacement has a component parallel to the scattering vector Q, the respective mode contributes to the structure factor. In a very pictorial

view you may regard the scattering process as a 'kick' in the direction of the momentum transfer. In the case of phonon creation, the neutrons 'kicks' off the phonon, in the case of annihilation the neutron 'gets kicked' by the phonon. If we compare the different scattering triangles in Fig. 4 b) we see, that in the upper triangle the q is nearly perpendicular to Q. Accordingly we measure at this position a transverse phonon branch, where the displacement is perpendicular to the wave vector. In the lower scattering triangle q and Q are parallel. Hence the neutron can only be scattered by a longitudinal phonon.
The intensity of a phonon mode depends inversely on its energy (eq. 17). Thus high energy phonons become difficult to measure. Expression 18 describes the occupation of a phonon mode. The $+$ sign refers to a phonon creation, the $-$ sign to annihilation.

$$n(\omega_j(\mathrm{q})) = (\exp[\hbar\omega_j(\mathrm{q})/k_bT] - 1)^{-1} \tag{20}$$

is the familiar Bose factor.

As eq. (16) contains the coherent nuclear scattering length that cross section describes the pair correlations in a crystal. Scattering provides also access to purely local dynamics of a scatterer via the incoherent scattering length, which comes from the deviations of the scattering length from different isotopes or different orientation of the neutron spin to the nuclear spin from the mean value. This incoherent scattering is described by the incoherent double differential cross section [14]:

$$\left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}\omega}\right)^{\pm}_{inc} = \frac{\mathrm{k}'}{\mathrm{k}}\frac{(2\pi)^3}{v_0} \tag{21}$$

$$\sum_i \frac{1}{2M_i}(\overline{b_i^2} - \overline{b_i}^2)\exp(-2W_i(\mathbf{Q})) \tag{22}$$

$$\times \sum_j \frac{|\mathbf{Q}\cdot\mathbf{e}_{ij}|^2}{\omega_j}\cdot\left(n(\omega_j(\mathrm{q})) + \frac{1}{2} \pm \frac{1}{2}\right)\delta(\omega \mp \omega_s) \tag{23}$$

From the incoherent scattering the phonon density of states can be derived, as it integrates over all possible phonon branches for a given energy.

4.2 Magnetic excitations

While in principle x-ray as a electro-magnetic wave could be scattered by the magnetic moments of a sample, the interaction away from the absorption edges is weak for the scattering to be detected by the present instruments. We will later come to the resonant inelastic scattering, which is a higher order scattering process which can be sensitive to magnetic excitations, too.
In contrast neutrons interact with the magnetic moments in the sample with a comparable strength as with the nuclei. Therefore inelastic neutron scattering is still the standard method to study magnetic excitations. The double differential cross sections for single magnon scattering

reads

$$\left(\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}\omega}\right)^{\pm}_{mag} = \frac{\mathrm{k}'}{\mathrm{k}}\frac{(2\pi)^3}{v_0} \tag{24}$$

$$\frac{1}{2}S\left(\frac{\gamma r_0}{2}gf(\mathrm{Q})\right)^2 \tag{25}$$

$$\exp\left[-2W(\mathrm{Q})\right] \tag{26}$$

$$(1+\hat{\mathrm{Q}}_z^2) \tag{27}$$

$$\sum_{\tau,\mathrm{q}}\left(n(\omega(\mathrm{q}))+\frac{1}{2}\pm\frac{1}{2}\right) \tag{28}$$

$$\delta(\omega\mp\omega(\mathrm{q}))\delta(\mathrm{Q}\mp\mathrm{q}-\tau). \tag{29}$$

This cross section has several distinct differences from the cross section for single phonon scattering eq. (16). The interaction strength is given by expression (25). It depends on the spin S or more general the total angular momentum J, the gyromagnetic factor of the neutron γ, the classical electron radius r_0, the Lande‘ factor g and the magnetic form factor $f(\mathrm{Q})$. The form factor reflects the extended distribution of the magnetic moment around the atom, as the outer atomic shells contain the unpaired electrons responsible for the magnetic moment. The expression (27) comes from the selection rule for the magnetic neutron scattering, which states that only the magnetic moment perpendicular to the momentum transfer scatters the neutron, and from the fact that the 1 magnon scattering reduces the magnetic moment in the quantization direction z and increases the x, y components, respectively.

4.3 Resonant Inelastic X-ray Scattering (RIXS)

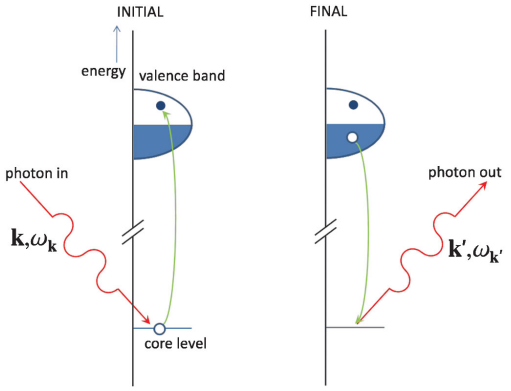
The cross sections discussed so far provide direct information about the dynamics of the lattice and the spin system and are therefore a solid test for any model describing the spin system. The coupling to the electron system can be investigated only indirectly by studying e.g. the life time of the excitations [15, 16]. In RIXS the opposite way is used: one excites the electron system and investigates, how the excitation couples to the different elementary excitations in the solid [17]. Here we have to distinguish two different processes: direct RIXS and indirect RIXS (Fig. 11).

In the direct process the core electron is excited into an empty valence band state and the core hole is filled by another valence band state. For this process the selection rules must allow both transitions. The direct process is e.g. allowed for the L edges of the transition metals in the soft x-ray regime. They excitations probed by this regime include single magnon excitations [18].

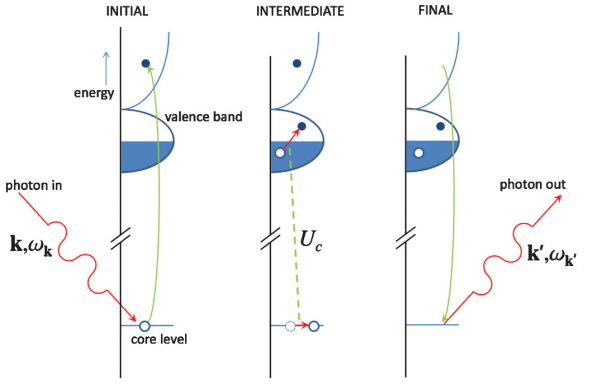
They indirect RIXS process is observed when the core electron is excited into an empty state in the conduction band relatively high above the Fermi energy. The core hole is then screened only weakly and acts as a scattering potential for the valence electrons. Alternatively one could say that the exciton of core hole and conduction band electron scatters off the valence band excitations. After the core hole is filled an electron-hole excitation is left behind in the valence band.

The local nature of the excitation into the intermediate state makes RIXS an element specific technique. It can therefore be used to study the dynamics in the subsystem of a complex material, e.g. the transition metal or the ligand in a transition metal oxide.

In principle RIXS is sensitive to a variety of elementary excitations e.g., plasmons, charge transfer excitations, crystal field excitations, magnons and phonons and more. The sensitivity states



(a) Direct RIXS process: The core electron is excited into an empty valence band state and the core hole is filled by a different valence band state. For both transitions to the core states must be allowed



(b) Indirect RIXS process: The core electron is excited far above the Fermi energy. The core hole potential scatters then the electrons in the valence band. After the reemission of the photon, this electron-hole excitation is left behind.

Fig. 11: *Schematic of the RIXS process*

also a problem since it is not always obvious, which excitation creates the RIXS signal. Additional interaction such as spin-orbit coupling for magnons are necessary because the photon field excites the electronic system of the sample.

Experimentally RIXS is today a medium resolution technique with a energy resolution $\Delta\hbar\omega \approx 0.1$ eV. This is due to the fact that the analyzers are not optimized for the final photon energy, which is set by the absorption edge of the element under study. The fourth generation x-ray sources such as energy recovery linac or XFEL provide bright prospects for the evolution of RIXS due to the again dramatically improved source brightness.

5 Conclusion

If one investigates the dynamics of a bulky sample on an atomic scale one has to rely on inelastic scattering methods using neutrons or x-ray photons. Both probes provide specific advantages, which must be weighted before proposing an experiment.
Inelastic neutron scattering provides an energy resolution down to several μeV. If you are interested in life times of the excitations, spin echo methods can even provide higher resolution. The double differential cross section is very well understood for phonon and magnon scattering and hence inelastic scattering may be the ultimate test for models of the dynamic in complex materials. The general problem of inelastic neutron scattering is the limited brightness of todays neutron sources requiring usually several grams of sample to be studied. This limits also the application of complex sample environments such as high pressure cells. Here the new intense spallation source increase significantly the parameter space that can be investigated.
IXS on the other hand is very well suited for the investigation of small sample volumes as the beam already has μm dimensions. As the photon energy is 6 to 8 orders of magnitude larger than the required energy resolution the monochromator development is very demanding as it faces the high energy load of the intense incoming beam. The reciprocal space region that can be probed is basically unlimited, which is very important for high energy excitations or excita-

tions at small momentum transfer e.g. in amorphous materials.
The emerging field of RIXS widens the field of elementary excitations that can be probed to electronic excitations. It can be used to study the coupling between the electrons and the phonon or magnons directly. Being element specific it can be used to distinguish between the ligand or metal dynamics in complex materials and how this affects the functional properties.

Both x-ray and neutron techniques face the challenge that cutting edge instruments are needed to address todays scientific challenges. It is therefore important to understand how these instruments work. Then one may set them up to push the limits further. In combination with the improvement of the sources, which will be provided by the further development of the facilities, novel insight in condensed matter research can be expected.

References

[1] B. Dorner. *Coherent Inelastic Neutron Scattering in Lattice Dynamics*. Springer-Verlag, Berlin, Heidelberg, New York, 1982.

[2] M. J. Cooper and R. Nathans. *Act. Cryst.*, 23:357, 1967.

[3] M. Popovici. *Act. cryst. A*, 31(JUL1):507 – 513, 1975.

[4] K. Lefmann, Ch. Niedermayer, A.B. Abrahamsen, C.R.H. Bahl, N.B. Christensen, H.S. Jacobsen, T.L. Larsen, P. Häfliger, U. Filges, and H.M. Rønnow. *Physica B: Condensed Matter*, 385-386, Part 2(0):1083 – 1085, 2006.

[5] M. Kempa, B. Janousova, J. Saroun, P. Flores, M. Boehm, F. Demmel, and J. Kulda. *Physica B: Condensed Matter*, 385-386, Part 2(0):1080 – 1082, 2006.

[6] A. Hiess, M. Jiménez-Ruiz, P. Courtois, R. Currat, J. Kulda, and F.J. Bermejo. *Physica B: Condensed Matter*, 385-386, Part 2(0):1077 – 1079, 2006.

[7] Wolfgang Schmidt and Michael Ohl. *Physica B: Condensed Matter*, 385-386, Part 2(0):1073 – 1076, 2006.

[8] J. Serrano, A. Bosak, M. Krisch, F. J. Manjón, A. H. Romero, N. Garro, X. Wang, A. Yoshikawa, and M. Kuball. *Physical Review Letters*, 106(20):205501, May 2011.

[9] J. Maultzsch, S. Reich, C. Thomsen, H. Requardt, and P. Ordejon. *Physical Review Letters*, 92(7):075501, 2004.

[10] T. S. Toellner, A. Alatas, and A. H. Said. *Journal of Synchrotron Radiation*, 18(4):605–611, May 2011.

[11] M RENNINGER. doppelspektrometer-auflosungsvermogens. *Zeitschrift Fur Naturforschung Part A-Astrophysik Physik Und Physikalische Chemie*, A 16(10):1110–&, 1961.

[12] Léon Van Hove. *Phys. Rev.*, 95(1):249–262, Jul 1954.

[13] Gen Shirane, Stephen M. Shapiro, and John M. Tranquada. *Neutron scattering with a triple axis spectrometer*. Cambridge University Press, Cambridge, 2002.

[14] G. l. Squires. *Introduction to the theory of thermal neutron scattering*. Cambridge university press, Cambridge, 1978.

[15] S. P. Bayrakci, T. Keller, K. Habicht, and B. Keimer. *Science*, 312(5782):1926–1929, 2006.

[16] T. Keller, P. Aynajian, K. Habicht, L. Boeri, S. K. Bose, and B. Keimer. *Physical Review Letters*, 96(22):225501, 2006.

[17] Luuk J. P. Ament, Michel van Veenendaal, Thomas P. Devereaux, John P. Hill, and Jeroen van den Brink. *Reviews of Modern Physics*, 83(2):705–767, June 2011.

[18] L. Braicovich, M. Moretti Sala, L. J. P. Ament, V. Bisogni, M. Minola, G. Balestrino, D. Di Castro, G. M. De Luca, M. Salluzzo, G. Ghiringhelli, and J. van den Brink. *Physical Review B*, 81(17):174533, May 2010.

D 5 Atomic Pair Distribution Function: Local Structure Investigations [1]

S. Disch
Laboratoire de Chimie des Polymères
Université Libre de Bruxelles, Belgium

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

With the emerging advances in nanotechnology, increasingly complex and nanostructured materials are explored for potential technological applications. These include non crystalline materials such as glasses, nanostructured materials such as nanoparticles, nanocomposites, or mesoporous compounds, and disordered crystalline materials, where the local structure deviates from the long range structure. In all these classes of materials, the physical properties are dominated by local structural correlations, and a precise description of this local structure is thus important for understanding the structure-property relationships. In most cases, quantitative structural analysis in an intermediate real space length scale of several nm is required. Local probes such as EXAFS and NMR spectroscopy are element specific and access only the very local environment of the first 2-3 coordination shells. Traditional crystallography, on the other hand, is based on periodic boundaries and translational symmetries and gives information on the long range average structure of the material under study. For amorphous or locally disordered materials, these requirements of periodic boundaries are not fulfilled anymore. This results in tremendous reflection broadening as well as diffuse scattering contributions which are usually discarded as scattering background in traditional crystallography. For description of the local structure, however, the diffuse scattering contribution is crucial. Based on Fourier transformation of the total scattering intensity including both Bragg and diffuse scattering contributions, analysis of the atomic Pair Distribution Function (PDF) gives access to both the local structural information and the average long range structure in real space.

Analysis of total scattering and the PDF has a long tradition for disordered materials, with early work on liquids and glasses such as liquid mercury [1], liquid sodium [2], and vitreous SiO_2 and B_2O_3 [3] carried out in the 1930s. With the advancement in computing as well as modern synchrotron radiation and spallation neutron sources, the technique is nowadays widely applied to nanostructured and locally disordered materials [4, 5, 6]. Data acquisition with improved resolution and a large range of the wavevector Q allows for studying the local structure of materials with with sharp features and in a large real space range as required for the study of nanomaterials. Ultrafast data acquisition allows for in-situ and time resolved studies [7, 8]. Automated data reduction as well as advanced modelling and refinement of local (dis)order and the PDF are facilitated by the availability of user friendly software.

This chapter provides a brief introduction to PDF analysis and its applications. It is based on a number of very illustrative and comprehensive reviews and textbooks on this topic [4, 5, 6, 9, 10, 11], which also provide further detailed reference. The chapter is organized as follows: A general introduction of the capabilities of the atomic pair distribution function is given in section 2, followed by the technical requirements for data acquisition, reduction, and analysis in section 3. In section 4, examples for application of the PDF method to a variety of materials are given in order to highlight the versatile use of this technique.

2 Total scattering and the atomic pair distribution function

Local atomic correlations appear as diffuse scattering in diffraction intensity profiles. The strong reflection broadening and diffuse scattering by nanoparticles as compared to the bulk material are shown in Figure 1a. In Rietveld refinements information on the local structure is lost, as the diffuse scattering and other background contributions are discarded. The total scattering structure function S(Q) contains both Bragg scattering and diffuse scattering contri-

butions and can be derived from the measured scattering intensity by a number of corrections and normalizations according to

$$S(Q) = \frac{I^{coh}(Q) - \sum c_i \mid f_i(Q) \mid^2}{\mid \sum c_i f_i(Q) \mid^2} \tag{1}$$

where $I^{coh}(Q)$ is the coherent scattering intensity from a powder sample after correction for background and experimental contributions and normalization to flux and number of atoms in the sample. Q is the magnitude of the wave vector $Q = 4\pi \sin(\theta)/\lambda$ with 2θ the scattering angle and λ the used radiation wavelength. c_i is the atomic concentration and $f_i(Q)$ the atomic form factor (in case of neutron scattering to be replaced by the neutron scattering length b_i) of the element i in the sample.

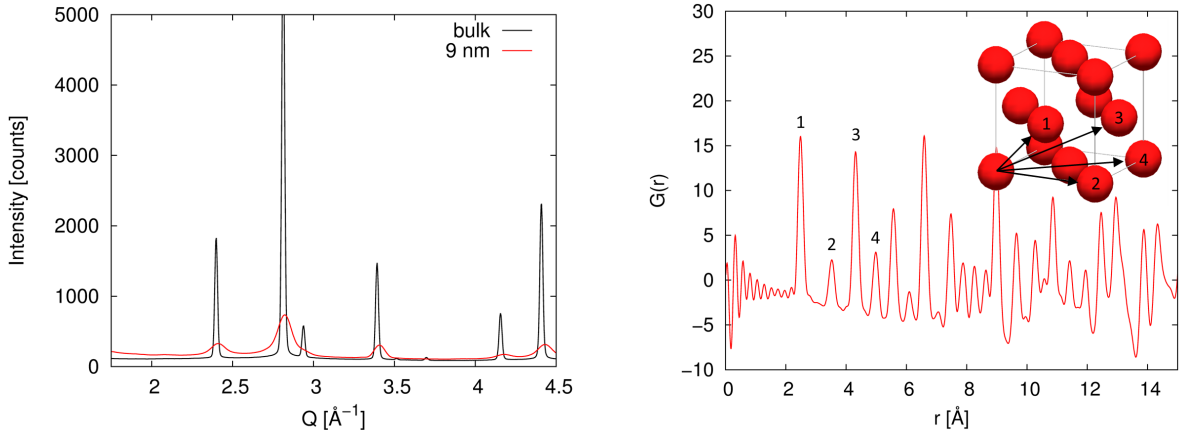


Fig. 1: a) Comparison of X-ray diffraction data for bulk magnetite and iron oxide nanoparticles. b) PDF obtained for Nickel powder. The first nearest neighbor correlations are indicated in the *fcc* crystal structure.

The $S(Q)$ is often presented as the reduced total scattering function $F(Q)$

$$F(Q) = Q[S(Q) - 1] \tag{2}$$

which emphasizes the scattering contributions at high Q. An example for $F(Q)$ is given in Figure 2c on page 7. The atomic pair distribution function (PDF) $G(r)$ is directly accessible from the $F(Q)$ through Fourier transform

$$G(r) = 2/\pi \int_0^\infty Q[S(Q) - 1] \sin(Qr) dQ \tag{3}$$

As the experimentally derived $S(Q)$ and $F(Q)$ include Bragg scattering as well as diffuse scattering contributions, the obtained PDF contains information on both the average and the local structure. The PDF is a continuous function reflecting all atomic distances within the material. The physical meaning of the $G(r)$ is a weighted probability of finding atomic pairs separated by r. It is defined as

$$G(r) = 4\pi r[\rho(r) - \rho_0] \tag{4}$$

where $\rho(r)$ is the microscopic pair density, ρ_0 is the average number density, and r is the radial distance. $G(r)$ approaches $-4\pi r\rho_0$ in the low r range and oscillates around zero for large r with positive peaks indicating atomic densities exceeding the average number density, whereas

negative peaks can be related to distances with less than average number density (note that due to negative neutron scattering lengths, atomic pair distances may also be indicated by negative peaks in neutron PDF). An example $G(r)$ for Nickel powder is given in Figure 1b.

Several distribution functions are distinguished in the literature, which are defined with different weighting and normalizations, but generally contain the same information, providing a direct distribution of atomic pair distances in real space [11]. However, for PDF analysis, the reduced $G(r)$ as defined in equation (4) is preferably used. The main reason for this is that $G(r)$ is directly obtained by Fourier transformation of the reduced total scattering function (eq. (3)) without any assumption of the average number density of the sample ρ_0. In contrast, the number density can experimentally be derived by analysis of the slope $-4\pi\rho_0$ of $G(r)$ at low r. Furthermore, after Fourier transform the uncertainties of $G(r)$ are constant in r, which facilitates refinement of a structure model and comparison to the obtained PDF.

The PDF allows for a direct determination of bond lengths from the peak positions. These bond lengths represent the average distance between the positions of pairs of atoms, as opposed to the distance between the average positions of atoms accessible by traditional crystallographic analysis [5]. This difference may become important for largely disordered systems or systems with pronounced lattice dynamics. Information on the bond length distribution due to lattice dynamics, correlated motion, or static disorder is gained from the peak shape in the PDF. A larger peak width indicates a broad distribution of bond lengths, whereas a narrow peak width corresponds to a sharp distribution of bond lengths, which is indicative for correlated motion. Furthermore, relative coordination numbers can be extracted in a quantitative analysis of the PDF from the integrated peak intensity. These may be known for crystalline materials from traditional crystallographic analysis, but in case of compositional disorder, the coordination numbers may reflect local correlations between site occupancies which are not accessible by traditional crystallography [5].

The real space resolution of the PDF is determined by the Q range of the total scattering function $S(Q)$ available for Fourier transformation. Since $S(Q)$ is obtained from experimental data, only a finite Q range can be probed. The termination of the $S(Q)$ at a suitable Q_{max} is important for generation of a reliable $G(r)$ because termination ripples will occur depending on the chosen Q_{max} as a result of the Fourier transformation. Termination ripples will be less pronounced and the r resolution obtained in the PDF will be better for a higher Q_{max}. On the other hand, if a large Q_{max} is chosen, low statistics in the high Q range may result in noise ripples in the obtained PDF. Whereas termination ripples can be modeled during refinement of the PDF, noise ripples are purely statistical and cannot be modeled and should thus be avoided. For this reason, a high Q_{max} is desired for generation of the PDF, but good statistics in the wide Q range is as important. For the study of nanomaterials, a good Q resolution is furthermore required in order to achieve a large r range in the obtained PDF. This is particularly useful for investigation of space resolved local structure deviations or the structural coherence length in nanosized materials. The structural coherence of the material under study is determined from the decay of the peak intensity at large r. For long range ordered crystalline materials, the $G(r)$ should exhibit peaks in an infinite r range, whereas for purely local correlations no PDF intensity is observed beyond the real space correlation length. Similar to reciprocal space analysis, where the peak width results from convolution of particle size (and lattice strain) broadening with the instrumental resolution, the damping of the PDF at large r contains contributions from both the structural coherence length and the instrumental resolution.

3 Data acquisition and analysis

3.1 Data acquisition

The focus of this chapter is on PDF analysis of microscopically isotropic samples or powders of crystalline materials. However, there exist examples for PDF analysis on oriented materials and single crystals of disordered materials in the literature, and further information on higher dimensional PDF can be found in [11]. In most cases, the experimental setup for measurements of the total scattering function $S(Q)$ is very similar to conventional powder diffraction experiments.

Samples for PDF analysis usually consist of disordered fine crystallites or nanocrystals. For synchrotron X-ray scattering, capillaries or flat plate transmission sample holders are used with a sample thickness of 0.5 - 1 mm. The thickness is optimal if the sample has an absorption coefficient of $\sim$ 50% at the chosen X-ray energy. However, in order to maintain a reasonable Q resolution, sample thicknesses should not exceed few mm. For neutron scattering experiments, cylindrical cans made of vanadium are used, and a larger amount of sample up to 10 g is required.

The most important requirements for total scattering measurements exceeding those for conventional powder diffraction are a wide accessible Q range with good statistics and an accurate separation of instrumental and sample related contributions to the $S(Q)$. As valid for all scattering techniques, a balance between good resolution in Q and high scattering intensity has to be found. Good Q resolution is mainly important for investigation of structural coherence lengths or r dependent parameters. If the sample should be studied using both PDF analysis and conventional crystallographic methods, such as Rietveld refinement, enhanced Q resolution is beneficial. According to $Q = 4\pi \sin(\theta)/\lambda$, a combination of large scattering angles and small wavelength (high energy) radiation is needed to achieve a large Q. With laboratory X-ray sources, a Q_{max} in the order of 8 Å^{-1} (Cu Kα) or 16 Å^{-1} (Mo Kα) can be achieved, which may be sufficient for studies of liquids and glasses with a low r resolution. However, for more ordered materials a higher r resolution is useful, and a wide Q range of 30 - 50 Å^{-1} is desired. Such a Q range is achieved by high-energy X-rays of 100 keV and higher at third generation synchrotron sources. A number of suitable beamlines at third generation synchrotron sources worldwide exists, with dedicated PDF beamlines including 11-ID-B at the Argonne Photon Source (APS), ID31 at the European Synchrotron Radiation Facility (ESRF), and P07 at PETRA III (Hamburg), just to name a few. The Q dependent atomic form factor decreases the coherent X-ray scattering intensity at large Q, resulting in apparently low signal in the high Q region of the diffraction pattern. However, the high X-ray flux at modern synchrotron sources compensates for the low scattering intensity at high Q by good statistics achievable in a reasonable amount of time. Good scattering statistics in a wide angular range in synchrotron X-ray experiments are achieved by using area detectors, which reduce the data collection time to seconds or minutes [7]. Ultrafast data acquisition allows for time-resolved measurements or in-situ measurements using different temperatures or gas environment [8, 12]. Energy-dispersive detectors are used to separate the elastic scattering intensity from inelastic Compton scattering, which is especially pronounced at large Q. For quantitative experiments, an accurate and stable beam monitor is required for normalization. Widely used beam monitors include ion chambers placed in the direct beam or scintillation detectors placed perpendicular to the beam in front of the sample. Normalization of the scattering intensity to absolute units is important if absolute peak intensities in the PDF are desired, such as for determination of coordination numbers.

Application of neutron scattering for PDF analysis is particularly useful in case of light scattering elements (such as H) in the sample or in case of a limited X-ray scattering contrast compared to an enhanced contrast in neutron scattering lengths. Neutrons furthermore have the advantage of Q independent scattering lengths allowing for good statistics at large Q. Time-of-flight (TOF) instruments at spallation sources are best suited for neutron PDF measurements because of the large flux of epithermal neutrons. Epithermal neutrons are undermoderated and thus high in energy, yielding high Q scattering data up to 100 Å^{-1}. Optimized neutron scattering instruments for PDF analysis include NPDF at Los Alamos National Laboratory (LANL), GEM at ISIS, and the new NOMAD instrument at the Spallation Neutron Source (SNS) in Oak Ridge National Laboratory (ORNL). In a typical neutron TOF diffraction instrument, the sample container is illuminated by a collimated neutron beam of about 1 x 4 cm. The sample space is surrounded by fixed detectors. Each detector element detects the full diffraction pattern as a function of the neutron time of flight. A gain in efficiency is achieved by covering a large solid angle with detectors, whereas detectors with the same scattering angles are grouped into detector banks. As a beam monitor, a thin vanadium foil can be placed in the beam. Detection of its incoherent scattering allows for monitoring both the incident flux and the incident neutron spectrum.
In both neutron and X-ray PDF measurements, low instrumental backgrounds are desired. Background scattering can be minimized by good collimation and shielding. In order to separate the purely coherent scattering from background contributions, reference measurements are made using the empty sample holder. For neutron PDF measurements, a vanadium absorber in the same sample holder is used for additional reference.

3.2 Data reduction

In a PDF experiment, the measured data is usually obtained in the form of intensity (normalized by monitor counts) as a function of the scattering angle 2θ or the time of flight and detector position, which is converted into the scattering vector Q. In order to derive the total scattering function $S(Q)$ and finally the PDF, several corrections and normalizations have to be carried out. An example for the data reduction process from the 2D area diffraction data via $I(Q)$ and $F(Q)$ to the PDF is given in Figure 2. Software for data correction and normalization as well as conversion to the PDF is available with PDFgetX2 [13] for X-ray measurements and PDFgetN [14] for neutron measurements.
First, the instrumental scattering contribution has to be separated from the measured data in order to obtain the pure sample-related scattering intensity $I(Q)$ (Figure 2b). This is usually done by measuring a number of background measurements.
Instrumental scattering contributions may arise from either the sample holder or the empty instrument (including air scattering). Due to absorption of the sample, the sample holder contribution may be overestimated from a single reference measurement. Reference measurements of both the sample holder and the empty instrument are thus required. In neutron scattering experiments, an additional vanadium reference has to be measured in order to account for the source spectrum.
Corrections which are carried out for both X-ray and neutron scattering experiments include the detector deadtime correction, absorption, and multiple scattering. Multiple scattering depends on the sample thickness and transparency, and is rather complicated to calculate. It may be neglected for thin samples in X-ray scattering experiments, but is relevant for neutron scattering experiments, where the sample absorption cross section is lower than the scattering cross section. Attenuation of the beam by a sample of known composition and density can be de-

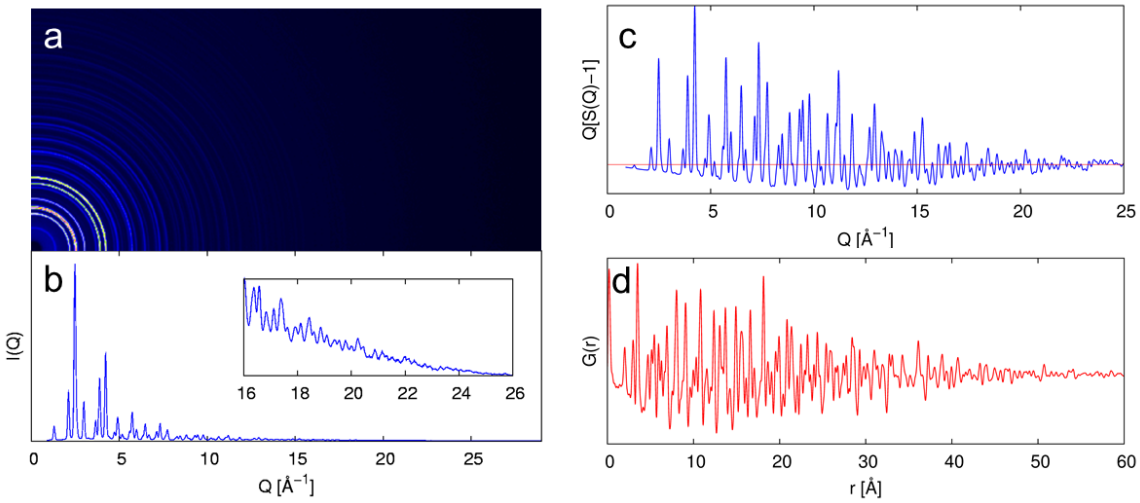


Fig. 2: Data reduction of high-energy X-ray scattering data to the PDF of magnetite: a) 2D scattering intensity. b) Radially integrated and normalized intensity $I(Q)$ after background subtraction. c) Reduced total scattering function $F(Q)$. d) Pair distribution function $G(r)$.

termined from tabulated values. A more precise determination of the experimental attenuation cross section of the sample is achieved by using beam monitors before and after the sample position.

Further important corrections for X-ray scattering experiments include the X-ray polarization and fluorescence as well as Compton scattering contributions. The X-ray polarization is relevant because X-rays become polarized upon scattering, which affects the angle dependence of the measured scattering intensity. This correction will be small for synchrotron experiments, where the beam is nearly 100% polarized, but important for experiments using laboratory sources. Fluorescent scattering can be reduced if the chosen X-ray energy is lower than the absorption edge of the most strongly scattering atomic species in the sample. Compton scattering is the inelastic, incoherent scattering contribution resulting from recoil of unbound or weakly bound electrons during the scattering process. Being a minor contribution at low Q, its intensity at high Q can be much larger than the coherent elastic scattering intensity. At the same time, the energy difference between elastic and Compton scattering increases, so that a large part of the Compton scattering can be separated using energy dispersive detectors. In the low Q region, Compton scattering can not be resolved from the elastic scattering, so that it is calculated theoretically in this region.

Corrections carried out for neutron scattering data may include the removal of magnetic scattering contributions. However, magnetic scattering occurs mainly at low Q due to the rapidly decreasing magnetic form factor, so that it contributes only weakly to the PDF. A more important correction for neutron scattering data by liquids is the correction of inelastically scattered neutrons incorrectly binned in Q. This effect is most pronounced in liquids of low atomic mass and can be corrected using the Placzek approach.

A more detailed description of all corrections including further references can be found in [11].

After application of all corrections and normalizations, the total scattering function $S(Q)$ is obtained. $S(Q)$ approaches 1 at high Q, whereas $F(Q)$ oscillates around zero at high Q (Figure 2c). The PDF is obtained by Fourier transform of the data according to equation (3) (Figure 2d). As explained above, special care is required for decision of the range of data used for Fourier

transform. In essence, a large Q_{max}, desired for minimization of termination ripples, has to be balanced against the signal to noise ratio at large Q.
Even in case of careful correction of the measured data, the obtained $S(Q)$ may still be imperfect, as many of the applied corrections might not be exact. For example, the sample attenuation coefficient depends on the sample density, which may be significantly lower in powder than in a bulk solid. As the applied corrections mostly have a low Q dependence, their over- or underestimation will lead to long wavelength oscillations of the obtained $S(Q)$. These result in sharp peaks in the very low r range of the obtained PDF, mostly below 1 Å, where no structural information is expected.
In order to optimize the $G(r)$ against such miscorrections, PDFgetX2 and PDFgetN provide an ad-hoc $G(r)$ optimization approach which modifies $S(Q)$ with multiplicative and additive coefficients according to

$$S'(Q) = \alpha S(Q) + \beta \tag{5}$$

The coefficients α and β can be refined in order to optimize the $G(r)$ at low r (so that the $S(Q)$ at large Q asymptotically approaches 1). However, the obtained PDF should still be handled with care to avoid misinterpretation of very low r features. In general, refinement of a broader r range will lead to more stable results.

Similar to anomalous scattering experiments, where the contrast enhancement close to an element's absorption edge is used for separation of the pure scattering contribution of the probed element, partial and differential PDFs can be derived, which specifically contain or exclude selected atomic element distances. The required contrast variation for the probed elements can be achieved either by high energy X-ray diffraction measurements close to an absorption edge of the chosen element, or by neutron data from samples containing different isotopes of the chosen element, or by combination of X-ray and neutron experiments, if the difference in scattering lengths is sufficient. MIXSCAT is a user friendly software for separation of the partial PDF is freely available [15].

3.3 PDF analysis

In the most straightforward way, the obtained PDF can be analyzed by refinement of individual nearest neighbor peaks, yielding to bond lengths and coordination numbers. The r dependence of the peak width will give information on possible correlated motion. For a more detailed analysis of the local structure, a structural model is required. The PDF can be calculated directly from the distribution of bond lengths, weighted by the scattering powers of the contributing atoms according to

$$G(r) + 4\pi r\rho_0 = \frac{1}{r}\sum_{\nu}\sum_{\mu}\frac{f_\nu(0)f_\mu(0)}{\langle f(0)\rangle^2}\delta(r - r_{\nu\mu}) \tag{6}$$

with $f_i(0)$ the atomic form factor of the ith atom at Q = 0 (in case of neutron scattering to be replaced by the neutron scattering length b_i), and $r_{\nu\mu}$ the distance separating the νth and μth atoms, whereas the sums run over all atoms in the sample. For a structural model with periodic boundaries, the first sum runs over the atoms in the unit cell only, which reduces the computational power demand tremendously. In order to account for the termination of the $S(Q)$, the calculated PDF is convoluted with a termination function $S(r) = \sin(Q_{max}r)/r$.

There are several different approaches to refine experimental PDF depending on the complexity of the structural model, and a number of programs for refinement is available.

For a model with periodic boundaries and a small number of a few 100 atoms, a least-squares refinement similar to the traditional Rietveld technique is possible. A user friendly program for this type of refinement is PDFgui [16]. The program allows for refinement of lattice parameters, atomic displacement parameters, and atomic positions in a given unit cell or a supercell thereof along with parameters for correlated motion. The structural coherence length is modeled using an envelope function accounting for the damping of peak intensities at high r. For this kind of refinement it has to be taken into account that the resulting parameters are only valid for length scales corresponding to the fit range. This gives the opportunity of r dependent refinements, yielding structural parameters as a function of the real space length scale. At large r the parameters will describe the average structure, whereas the local structure is probed in the low r range.

The DISCUS program is capable of creating locally disordered structures in a parameterized approach and simulating both diffraction patterns including the diffuse scattering and the PDF [17]. Starting from a symmetric structure with few parameters, disorder can be introduced in the form of short-range order, modulations, domain structures, or stacking faults. For nano-sized materials, boundary conditions are assigned to the model before PDF calculation. By this approach, crystallography-constrained models of some 1000s of atoms are created using a small number of parameters. The program DIFFEV is part of the DISCUS distribution and provides refinement of the PDF using a differential evolutionary algorithm. It allows for a higher level modeling of a few model dependent parameters such as the particle size, stacking fault probability, or lattice parameter in order to match the experimental PDF including *e. g.* both structure and morphology of nanoparticles. A very detailed discussion of the modeling of disordered structured using DISCUS and DIFFEV is given in a recent textbook on defect structure simulations [18].

If a starting model with a large number of atoms (several 10000) is given, the Reverse Monte-Carlo (RMC) technique is a useful approach. The technique is based on a calculation of the PDF of the starting model and determination of the residuals function χ^2. In each simulation step, a random site is moved, and the improvement or not of the new χ^2 is assessed. The residuals function is minimized using a simulated annealing algorithm. Further details about the RMC technique applied to disordered crystalline materials is given by Keen *et al.* [19]. A program available for this refinement technique is the RMCProfile software [20]. It allows for refinement of the total scattering and integrated Bragg intensities, constraining the average structure of the model to the crystallographic result.

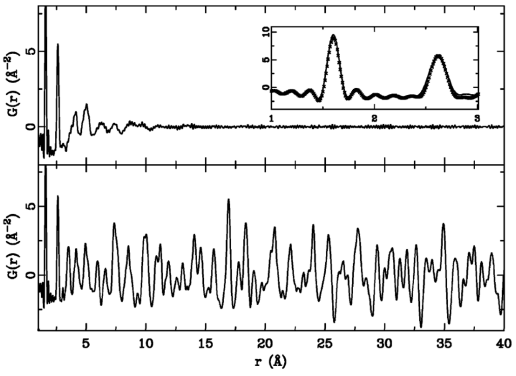
The first example for ab-initio determination of solid-state nanostructure from PDF data alone has been demonstrated for C_{60} fullerene using the LIGA algorithm [21]. This algorithm was developed and applied to the structure solution from precise unassigned interatomic distances only. In the future, more complex systems with larger clusters of lower symmetry may be accomplishable by incorporating information from complementary data.

Examples for application of some of these analysis techniques to a variation of disordered materials are given in the following section.

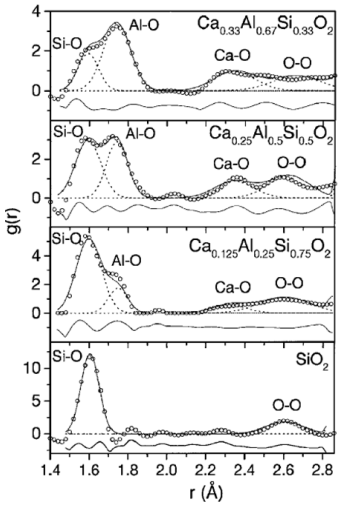
4 Application of PDF analysis

4.1 Liquids and Glasses

Liquids and glasses exhibit only very low range structural correlations, corresponding to the first few coordination spheres. Consequently, the PDF of liquids and glasses is characterized by very few first coordination peaks and a fast decay in peak intensity in increasing r. For comparison the PDFs obtained by a fused silica glass rod and a crystalline quartz powder are shown in Figure 3a [6]. Whereas the crystalline sample exhibits sharp peaks throughout the entire r range, the PDF of the glass sample decays very quickly after the first sharp peaks. These peaks correspond to the Si-O and Si-Si nearest neighbor distances, illustrating that the local environment is the same for silica glass and crystalline quartz.



(a) Comparison of the experimentally obtained PDF of silica glass (top) and crystalline quartz (bottom). Inset: direct comparison of the first nearest neighbor distances of both data sets [6]



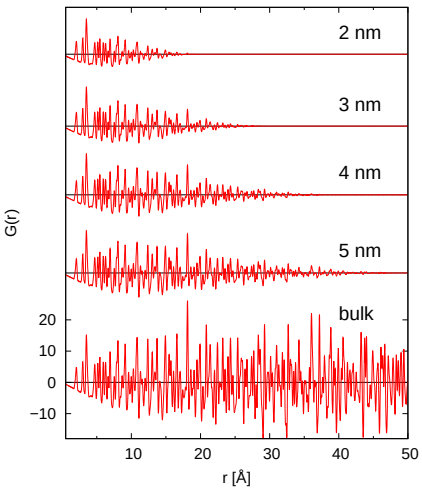
(b) Nearest neighbor distances in the PDF of calcium aluminosilicate glasses [22].

Fig. 3: Examples for PDF by silica glasses.

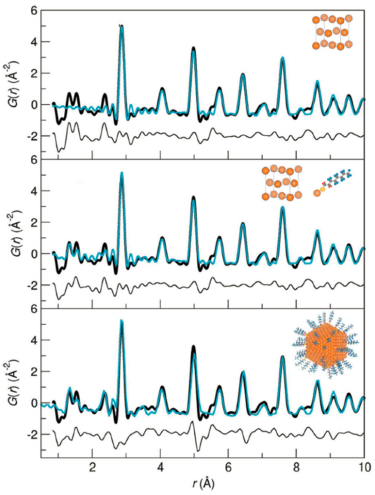
An illustrative example of the capability of the PDF technique in this field has been given by Petkov *et al.* in a study of the network connectivity in calcium aluminosilicate glasses [22]. Improved r resolution allowed to resolve the atomic distances of the SiO_4 and AlO_4 tetrahedra building up the glass network (Figure 3b). Determination of their coordination numbers in dependence of the Ca (and Al) concentration revealed a reduced coordination number of Si with increasing Al concentration. So-called nonbridging oxygens, which are responsible for breaking of the network with increasing Al concentration, were thus assigned selectively to the SiO_4 tetrahedra. In order to achieve the required r resolution, the total scattering function $S(Q)$ was determined precisely in a wide Q range up to 40 $Å^{-1}$. Improvement of the statistical accuracy of the experimental data in the high Q range was achieved by smoothing procedures. An important task was the extraction of the elastic component of the diffracted X-ray intensity, which was accomplished by use of an energy-dispersive detector and additional application of the Ruland window function for removal of residual Compton scattering at low Q.

4.2 Nanoparticles

Nanoparticles are intensely investigated in many fields of research as the reduction of size is usually accompanied by significant effects on their physical properties as compared to the respective bulk materials. Although nanoparticles might in a first approximation be regarded as nanocrystals with an atomic structure similar to the bulk, local structural effects become dominant with decreasing particle size. The most obvious effect of the particle size on the measured PDF is the smaller structural coherence length resulting in a massive damping of the PDF at large r as compared to the bulk material, see Figure 4a. The damping of the PDF can be described by an envelope function depending of the particle shape and size, which is multiplied with the calculated PDF for the bulk average structure and the instrumental damping factors [23]. Such envelope functions have been derived for spherical particles including particle size distributions and different particle shapes [23, 24]. The envelope function for monodisperse spherical particles is implemented in the PDFgui software [16]. The simulated PDF for nanoparticles with different particle size are shown in Figure 4a.



(a) Comparison of the PDF as simulated for bulk magnetite and nanoparticles of different diameter, illustrating the particle size effect on the PDF.



(b) Experimental PDF of Au nanoparticles with refinements of the nanoparticle structure as a single phase (top), the nanoparticle structure and an additional ligand phase (center), and the entire nanoparticle with ligand chains (bottom) [25]

Fig. 4: Examples for PDF by nanoparticles.

Page *et al.* have recently demonstrated the variation of the envelope function by different particle sizes and shapes through simulation of the PDF for entire nanoparticles utilizing DISCUS and DIFFEV [25]. By this approach, the entire nanoparticle structure is modeled (including the refinement of a phenomenological polynomial to match the baseline of the experimental PDF [25]) instead of calculating the PDF of an infinite structure multiplied with an analytical envelope function. It was further shown that such whole-particle structure modeling allows for differentiation of disorder of the nanoparticle compositional against core-shell structures and even provides insight into the surface coverage of an organic ligand shell. Whereas the presence of a thiol ligand shell in the studied gold nanoparticle samples has been confirmed by a

two phase refinement of the nanoparticle and the ligand phases, modeling of the entire particle including the ligand shell allowed for determination of the surface coverage and the Au-S distance (Figure 4b). Knowing the molecular structure of the ligand, the orientation of the ligand with respect to the particle radial direction was defined. Further refinement parameters include the Au lattice parameter, an atomic displacement parameter, and a cuboctahedral particle size.
Further examples of the DISCUS approach to modeling nanoparticle PDF include the determination of nanoparticle structure, shape, and defect structure in II-VI semiconductor nanoparticles [26]. Modeling of the PDF revealed an high stacking fault density between wurtzite and zincblende modifications which was not accessible by Rietveld refinement. For CdSe/ZnS core shell nanoparticles PDF analysis further revealed the structurally incoherent attachment of particle core and shell.

4.3 Local Disorder in Crystalline Materials

Even for crystalline materials, exhibiting long range order, disorder can be present on a local scale. In many cases this local disorder directly affects physical properties of the materials, which makes PDF analysis an important tool for understanding the structure-property relations. An illustrative example for this are the perovskite manganites, which are intensely investigated due to their colossal-magnetoresistance properties. Perovskite manganites such as $LaMnO_3$ consist of MnO_6 octahedra which are known to undergo Jahn-Teller distortion, resulting in a variation of Mn-O distances. Using PDF analysis, Qiu *et al.* have shown that even above the transition temperature, where traditional crystallography reveals almost regular MnO_6 octahedra, the distortion is preserved locally [27]. PDF analysis in different r ranges with increasing r_{max} revealed constantly distorted octahedra at low temperatures, where the Jahn-Teller distortion is observed by crystallographic analysis. Above the transition temperature, the Mn-O distances determined with a large r_{max} indicated perfect MnO_6 octahedra in agreement with crystallographic analysis. However, on the local scale, using a shorter real space range for refinement of the PDF, the determined Mn-O distances revealed distorted octahedra even at high temperatures (Figure 5). A study of the $La_{1-x}Ca_xMnO_3$ series by Božin *et al.* revealed the variation of the Mn-O distance with x and temperature [28]. It was shown that the variation of Mn-O distances determined in a local r range coincides well with the CMR transition temperature, a relation which was not apparent from the long range structure.
A different effect of local disorder in perovskite materials is the appearance of polar nanoregions in relaxor ferroelectrics. Despite the cubic symmetry of the bulk material, off-center displacements of the Pb atoms in the PbO_6 octahedra in $Pb(Mg_{1/3}Nb_{2/3})O_3$ and related compounds generate a local polarization which is correlated in polar domains. PDF analysis demonstrated the existence of local polarization and is sensitive to the volume fraction of these domains with dimensions of 5-50 Å [29]. Volume fractions of up to 30% have been observed in temperature dependent measurements.

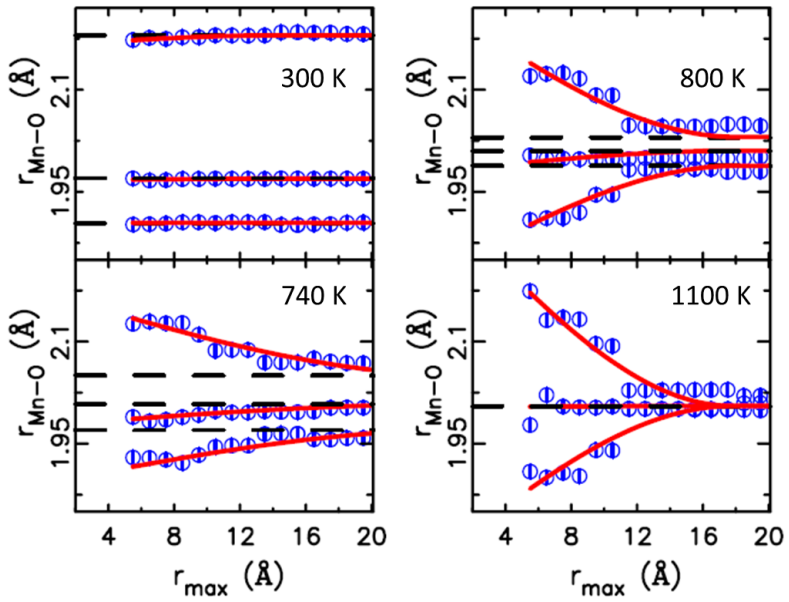


Fig. 5: Mn-O bond lengths in $LaMnO_3$ obtained from PDF at different temperatures and refined as a function of r_{max} [27]

References

[1] P. Debye, H. Menke, Physikal. Zeit. **31**, 797 (1930)

[2] L. P. Tarasov, B. E. Warren, J. Chem. Phys. **4**, 236 (1936)

[3] B. E. Warren, H. Krutter, O. Morningstar, J. Am. Ceram. Soc. **19**, 202 (1936)

[4] S. J. L. Billinge, M. G. Kanatzidis, Chem. Commun. **7**, 749 (2004)

[5] C. A. Young, A. L. Goodwin, J. Mater. Chem. **21**, 6464 (2011)

[6] T. Proffen, H. Kim, J. Mater. Chem. **19**, 5078 (2009)

[7] P. J. Chupas, X. Qiu, J. C. Hanson, P. L. Lee, C. P. Grey, S. J. L. Billinge, J. Appl. Cryst. **36**, 1342 (2003)

[8] P. J. Chupas, K. W. Chapman, P. L. Lee, J. Appl. Cryst. **40**, 463 (2007)

[9] S. J. L. Billinge J. Solid State Chem. **181**, 1695 (2008)

[10] V. Petkov, Mater. Today **11**, 28 (2008)

[11] T. Egami, S. J. L. Billinge, *Underneath the Bragg Peaks: Structural Analysis of Complex Materials* (Pergamon Press, Elsevier, Oxford, England, 2003)

[12] P. J. Chupas, K. W. Chapman, G. Jennings, P. L. Lee, C. P. Grey, J. Am. Chem. Soc. **129**, 13822 (2007)

[13] X. Qiu, J. W. Thompson, S. J. L. Billinge, J. Appl. Cryst. **37**, 678 (2004)

[14] P. F. Peterson, M. Gutmann, T. Proffen, S. J. L. Billinge, J. Appl. Cryst. **33**, 1192 (2000)

[15] C. Wurden, K. Page, A. Llobet, C. E. White, Th. Proffen, J. Appl. Cryst. **43**, 635 (2010)

[16] C. L. Farrow, P. Juhas, J. W. Liu, D. Bryndin, E. S. Bozin, J. Bloch, T. Proffen, S. J. L. Billinge, J. Phys.: Condens. Matter **19**, 335219 (2007)

[17] T. Proffen, R. B. Neder, J. Appl. Cryst. **32**, 572 (1999)

[18] R. B. Neder, T. Proffen, *Diffuse Scattering and Defect Structure Simulations* (Oxford University Press, New York, 2008)

[19] D. A. Keen, M. G. Tucker, M. T. Dove, J. Phys.: Condens. Matter **17**, S15 (2005)

[20] M. G. Tucker, D. A. Keen, M. T. Dove, A. L. Goodwin, Q. Hui, J. Phys.: Condens. Matter **19**, 335218 (2007)

[21] P. Juhas, D. M. Cherba, P. M. Duxbury, W. F. Punch, S. J. L. Billinge, Nature **440**, 655 (2006)

[22] V. Petkov, S. J. L. Billinge, S. D. Shastri, B. Himmel, Phys. Rev. Lett **85**, 3436 (2000)

[23] R. C. Howell, T. Proffen, S. D. Conradson, Phys. Rev. B **73**, 094107 (2006)

[24] K. Kodama, S. Iikubo, T. Taguchi, S. Shamoto, Acta Cryst. **A62**, 444 (2006)

[25] K. Page, T. C. Hood, T. Proffen, R. B. Neder, J. App. Cryst. **44**, 327 (2011)

[26] R. B. Neder, V. I. Korsunskiy, Ch. Chory, G. Müller, A. Hofmann, S. Dembski, Ch. Graf, R. Rühl, phys. stat. sol. (c) **4**, 3221 (2007)

[27] X. Qiu, Th. Proffen, J. F. Mitchell, S. J. L. Billinge, Phys. Rev. Lett. **94**, 177203 (2005)

[28] E. S. Božin, M. Schmidt, A. J. DeConinck, G. Paglia, J. F. Mitchell, T. Chatterji, P. G. Radaelli, Th. Proffen, S. J. L. Billinge, Phys. Rev. Lett. **98**, 137203 (2007)

[29] I. K. Jeong, T. W. Darling, J. K. Lee, Th. Proffen, R. H. Heffner, J. S. Park, K. S. Hong, W. Dmowski, T. Egami, Phys. Rev. Lett. **94**, 147602 (2005)

D6 Critical Scattering at Phase Transitions

K. Nemkovski
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction to critical phenomena

Critical phenomena are a fascinating field that has been a subject of intensive studies for several decades. There are many books and reviews devoted to the critical phase transitions and related problems. This lecture follows mainly to the famous book by M. F. Collins [1].

Let us start our consideration from the analysis of the phase diagram for the typical gas-liquid-solid system shown in Fig. 1. The system demonstrates three phases: vapor, liquid, and solid. There are areas where the system can be said to be in one of these three phases. These areas are separated by boundary lines known as *phase boundaries.* Two phases can coexist at phase boundaries; the system is found to be heterogeneous with regions corresponding to each of the phases. For instance, on the vapor-liquid phase boundary the system has part of its volume occupied by liquid and part by vapor with a surface separating the liquid and vapor. Usually gravity will have the effect of pulling the liquid phase to the bottom of the volume.

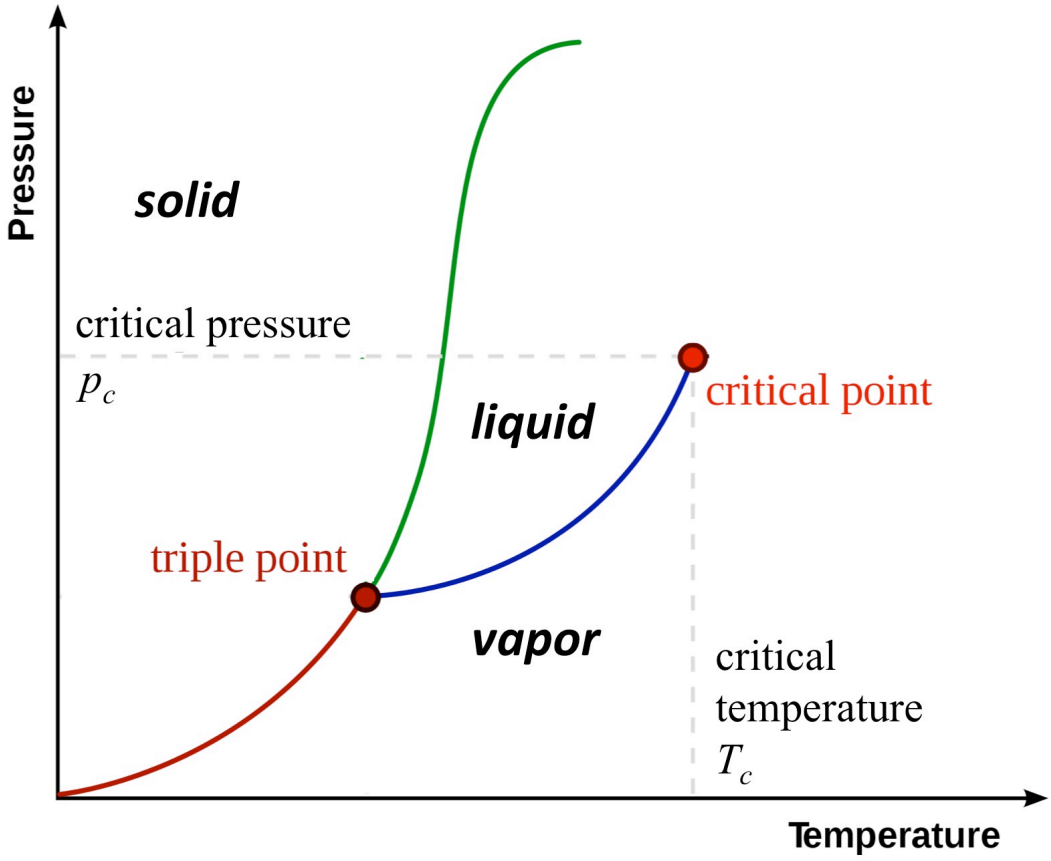


Fig. 1: *Pressure-temperature diagram for a typical gas-liquid-solid system (adapted from [2]).*

The point where the three phase boundary lines on Fig. 1 intersect is known as the *triple point.* All three phases can coexist at this particular temperature and pressure. Each phase will occupy its own local volume with surface boundaries separating one phase from another.
If the system is taken across any of the lines in Fig. 1, by varying the temperature or pressure, there is a discontinuous change of properties and a latent heat is exhibited. A remarkable feature of Fig. 1 is that the phase boundary between liquid and vapor simply vanishes out at a point known as the *critical point* (the corresponding temperature is called critical temperature and usually denoted as T_c). At this point the liquid and vapor become indistinguishable. By going round the critical point in Fig. 1 on the right-hand side it is possible to take a path from the region marked "vapor" to the region marked "liquid" without crossing any phase boundary. Thus, we can go from the vapor to the liquid without experiencing any discontinuous change in the properties.

These features can be described equivalently in terms of the energy of the system. The appropriate energy function for the variables p and T is the *Gibbs free energy G(P, T).* When two phases coexist at a phase boundary their free energies must be equal. However, there is no reason why the partial derivatives $\partial G/\partial p$ or $\partial G/\partial T$ need be the same at the boundaries and normally they would be different, so that in crossing a phase boundary there would be a discontinuity in the first derivative of G. Such a transition is called a *first-order phase transition* because the first derivative of G is discontinuous.

Phase transitions that pass through a critical point are known as *critical phase transitions.* Historically, such phase transitions were labeled as second-order phase transitions, in the expectation that they would correspond to a discontinuity in the second derivative of G. However, it became apparent later that this classification was not appropriate, because the behavior of G at a critical point is found to be non-analytic, so the phase transition was renamed as a critical phase transition.

First-order phase transitions are difficult to understand on a microscopic scale because the physical properties of the two phases are radically different. The entropy difference between the phases is sufficiently large in most cases to preclude statistical fluctuations in each phase that give short-lived micro-regions of the other phase close to the phase boundary and the phase transition may have to take place by nucleation of the second phase at some special points, such as impurities or boundary walls, followed by growth of the regions of the second phase.

In critical phase transitions, short-lived fluctuating micro-regions of one phase in the other are always found; they give rise to critical scattering, which is the subject of this lecture. As the critical point is approached, the size of these regions increases and becomes infinite at the critical point.

Critical phase transitions with the same general properties are also found for magnetism ferroelectricity, superconductivity, superfluidity in liquid helium, mixing of liquids, and ordering in alloys. It is quite amazing that these systems, which have very different microscopic Hamiltonians, all show similar types of phase transitions.

Below we list the properties that are common to all critical phase transitions [1].

1. There is a symmetry that is broken at the critical point. This symmetry is represented by

a parameter η, known as the *order parameter,* that is a continuous function of the temperature; η is zero for temperatures above T_c and is nonzero for temperatures below T_c. For ferromagnetism the order parameter is the magnetization, while for the liquid-gas transition it is the difference between the densities of the liquid and the gas phases.

The order parameter may be a scalar or a vector of dimensionality *D.* The dimensionality of the order parameter has a strong influence on the properties of the phase transition.

2. Near the critical point there are fluctuating micro-regions of both phases involved. The linear extent of these regions (the *correlation length,* ξ) tends to infinity as the critical point is approached from any direction.

3. The response time of the system tends to infinity as the critical point is approached from any direction. This is known as *critical slowing down.*

2 Elements of theory

2.1 Order parameter and critical exponents

The simplest treatment of continuous phase transitions is the Ginzburg-Landau theory (also sometimes called the mean-field theory or just the Landau theory). The basic assumption of this theory is that the thermodynamic energy functions can be expanded as Taylor series near the critical point. This assumption is in fact not correct, as the functions are not analytic at the critical point. However, the theory does hold for interactions of infinitely long range and for systems with four or more dimensions. Its simplicity and generality make it a good starting point for treatments of continuous phase transitions, and it allows us to introduce some concepts and frameworks that will be needed later.

Let us consider a phase transition from a ferromagnetic to a paramagnetic state. Then the order parameter is the magnetization and the energy function appropriate for the variables T and $\mathbf{M}$ is the Helmholtz free energy $F(T, \mathbf{M})$, or $F(T, \boldsymbol{\eta})$. Near the critical point the order parameter will be small and $F(T, \boldsymbol{\eta})$ can be expanded as:

$$F(T,\boldsymbol{\eta}) = F_0(T) + \alpha_2(T)\eta^2 + \alpha_4(T)\eta^4 + \ldots \qquad (1)\ ,$$

where $\eta^2 = \boldsymbol{\eta} \cdot \boldsymbol{\eta}$. Because the free energy is the same for a magnetization $\boldsymbol{\eta}$ and for its negative $-\boldsymbol{\eta}$, odd terms will not appear in the expansion.

The merit of the Ginzburg-Landau model is that it is solvable in just about every case. On the other hand, in many cases its predictions differ from the behavior observed experimentally. For instance, for the mentioned above transition between a ferromagnetic and a paramagnetic states the Ginzburg-Landau theory predicts a downward discontinuity in the specific heat at the critical temperature, as shown in Fig. 2:. In fact the specific heat diverges at T_C in what is usually referred to as a lambda anomaly; this is shown by the dashed line. Lambda anomalies correspond to rather weak divergences at the critical temperature, with a form similar to a logarithmic divergence. This result shows quite clearly that the Ginzburg-Landau theory is inadequate.

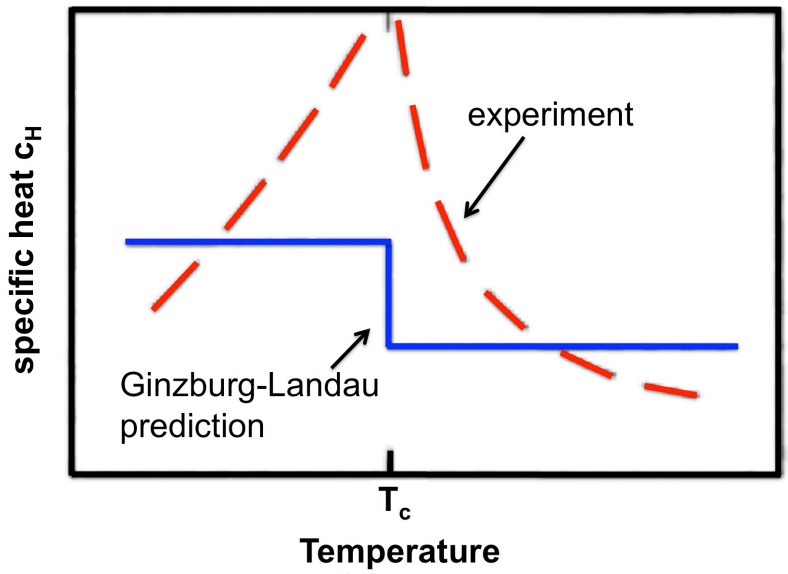


Fig. 2: *Variation of the specific heat C_H with temperature near a critical phase transition (adapted from [1]). The solid line shows the predictions of Ginzburg-Landau theory; the dashed line represents experimentally observed behavior of the specific heat.*

Magnetic susceptibility predicted by the Ginzburg-Landau theory follows to the Curie-Weiss law. It is found experimentally to be not too bad at temperatures well above T_c but to break down close to T_c. Thus, model derivation would predict just the opposite behavior, since a Taylor expansion about the critical point should be most accurate close to that point.

In fact there are two types of continuous phase transition for which the Ginzburg-Landau theory seems to be satisfactory. These are for ferroelectrics and for type I superconductors. The common feature of these two cases is that the microscopic interactions are of long range, and the Ginzburg-Landau theory actually gives correct predictions for interactions of infinitely long range. However, for most critical phase transitions the interactions are short-ranged and a better theory is needed.

When a theory fails, as does the Ginzburg-Landau theory, it is often appropriate to try to rationalize all the experimental data phenomenologically. In the magnetic case, for example, there are measurements available near the critical point of the susceptibility, the specific heat, the correlation length, and the magnetization. The first three of these properties all become very large as the critical point is approached. Experiment cannot give a result of infinity for physical quantities, but the "better" the conditions the larger is the value measured, and it is generally believed that these quantities are in fact becoming infinite at the critical point. In an analogous way the magnetization tends to zero as the critical point is approached, so that the reciprocal of the magnetization tends to infinity.

Many careful measurements have been made of how these parameters diverge as the critical point is approached, and almost all of them seem to obey a simple power law.

To give an example, if we define a reduced temperature t by $t = \frac{T - T_C}{T_C}$, then, close to the critical point, measurements of the isothermal susceptibility χ_T in small fields at temperatures just above T_C, fit the law

$$\chi_T = at^{-\gamma} \qquad (T > T_C) \qquad (2)\ ,$$

where a and γ are constants. This is illustrated in Fig. 3:, which shows measurements by Noakes et al. [3] of the susceptibility of iron alloyed with 0.5% tungsten plotted as a function of $T - T_c$. The plot is on a log-log scale and shows a straight line with slope –1.33, corresponding to $\gamma = 1.33$.

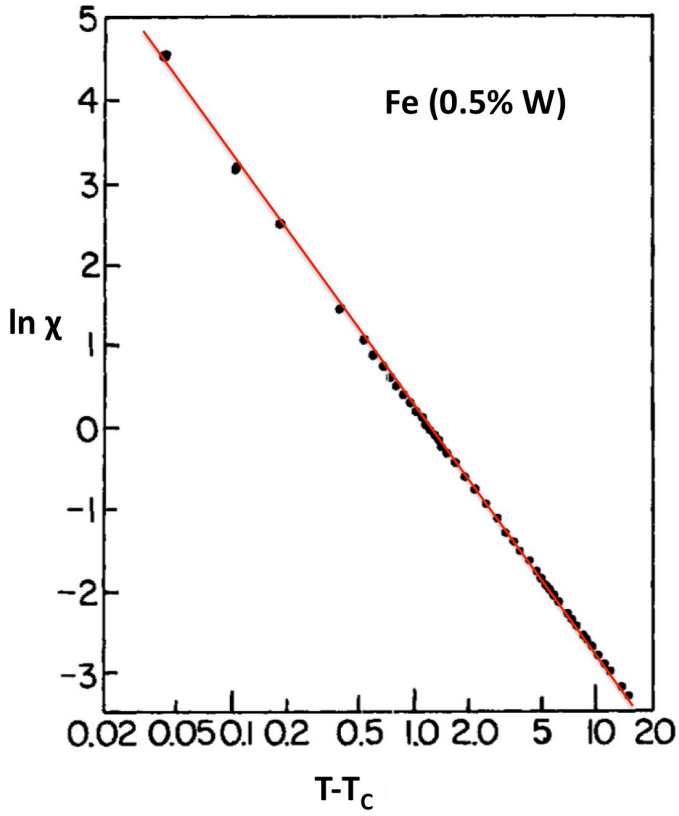


Fig. 3: *Log-log plot of the susceptibility of a ferromagnetic against $T - T_C$ (adapted from [3]). Solid line is a guide for the eye.*

The indices of the power laws, such as γ, are known as *critical exponents.* Their values vary somewhat from case to case, but the power law itself almost always holds. More carefully, one can say that a function $f(x)$ has critical exponent λ close to the critical point $x = 0$ as the critical point is approached from positive x, we mean $f(x) \sim x^{\lambda}$ when $x \to 0+$. The exponent λ can be defined by:

$$\lambda = \lim_{x \to 0+} \frac{\ln f(x)}{\ln x} \qquad (3)\ ,$$

so that λ is defined by the asymptotic behavior close to the critical point. If the critical point can be approached from negative x also, then we can analogously define a critical exponent λ' for $x \to 0-$. A summary of these exponents for different physical characteristics is given in Table 1.

Table 1. Definitions of some static magnetic critical exponents (adapted from [1]. and [4]).

Critical exponent	Property	Power law	Conditions
γ	Susceptibility, χ_T	$t^{-\gamma}$	$T > T_c$, H = 0
γ'	Susceptibility, χ_T	$(-t)^{-\gamma'}$	$T < T_c$, H = 0
β	Magnetization, M	$(-t)^{\beta}$	$T < T_c$, H = 0
δ	Magnetization, M	$H^{1/\delta}$	$T = T_c$
α	Specific heat, C_H	$t^{-\alpha}$	$T > T_c$, H = 0
α'	Specific heat, C_H	$(-t)^{-\alpha'}$	$T < T_c$, H = 0
ν	Correlation length, ξ	$t^{-\nu}$	$T > T_c$, H = 0
ν'	Correlation length, ξ	$(-t)^{-\nu'}$	$T < T_c$, H = 0
η	Pair correlation function $G(\mathbf{R})$	$\lvert\mathbf{R}\rvert^{2-d-\eta}$	$T = T_c$, H = 0 (d is dimensionality of the system)

The standard experimental practice in dealing with critical properties is to plot the logarithm of *f(x)* against the logarithm of x. An example of this was seen in Fig. 3:, where *f(x)* = χ *and* $x = T - T_c$, and these are plotted logarithmically. An immediate problem with making such a plot is that T_c must be known a priori, and this is often not the case. Uncertainties in T_C affect the plot very markedly at small x but have little effect at large x, and often the plot has to be made with T_c as a fitted parameter - fitted so as to keep the plot a straight line at the smallest values of $T - T_c$. Such a procedure clearly introduces experimental error into the

determination of critical exponents. It becomes hard to estimate the uncertainty in the critical exponent, since the slope of the plot at small x is sensitive to the fitted value of T_c and the slope at large x is sensitive to correction terms. In consequence, practical determinations of critical exponents usually have limited accuracy and very few measurements are able to determine critical exponents to much higher absolute accuracy than 0.005; most experimental data show uncertainties in critical exponents in the region 0.01-0.02 because of the difficulties that have been outlined [1].
We have also to mention that it is possible to derive from thermodynamic arguments certain exact relationships between response functions. For example, in the magnetic case it can be shown that

$$\alpha' + 2\beta + \gamma' > 2 \tag{4}$$

This is known as the *Rushbrooke inequality* [5]. More detailed discussion of the critical exponent inequalities can be found in the book [4].

2.2 Universality and standard models

A look at all the experimental data on critical exponents reveals certain apparently systematic trends. In the (relatively few) cases where critical exponents can be calculated with reasonable accuracy from theoretical models, the same features are apparent. These are usually expressed as the hypothesis of *universality* which states that [1]:

For a continuous phase transition the static critical exponents depend on the following three properties and nothing else:

1. The dimensionality of the system, *d.*
2. The dimensionality of the order parameter, *D* (or more precisely, the symmetry of the order parameter; in simple cases this is equivalent to the number of dimensions in which the order parameter is free to vary).
3. Whether the forces are of short or long range.

This represents an enormous generalization, because it implies that the nature of the microscopic interaction is irrelevant (except for its range, if that range is long). It is also irrelevant whether the system is quantum-mechanical or classical. For continuous transitions within the solid state, such as magnetic transitions, the critical exponents are predicted to be the same whatever the crystal structure.

In addition to the static properties covered by universality, there are also dynamic properties that show divergences at critical points that obey power laws with critical exponents. Universality does not work for these exponents. A simple example of the breakdown is provided by dynamics of three-dimensional Heisenberg systems. Ferromagnetics have long-wavelength excitations in the ordered state (spin waves) that have energy proportional to q^2, where q is the wave vector, while antiferromagnetics have long-wavelength excitations that have energy proportional to q. This results in different dynamic critical properties for systems that differ only in the sign of the coupling between spins. This difference can be expressed by the fact that for the ferromagnetic the order parameter (the magnetization) commutes with the

Hamiltonian and so is conserved, while for the antiferromagnetic the order parameter is the staggered magnetization, which does not commute with the Hamiltonian and is not conserved. It was proposed that universality holds for dynamic properties if a fourth condition is added to the three listed. This is [1]:

4. The dynamic critical exponents also depend on the conservation laws of the system.

Universality offers the prospect of a simplification in the theory of continuous phase transitions. There are many very different systems that show such phase transitions: for most of them the microscopic Hamiltonian is quite complicated and there is little hope of making direct calculations near the critical point. However, the universality enables us to choose the simplest theoretical model for any particular *universality class* (i.e., any particular values of *D, d,* and the range of forces). If we can solve this model, then the result will apply to all models within the universality class. These "simplest" models are magnetic, which is why critical phenomena are often discussed in the language of magnetism. The models are for different values of the dimensionality of the order parameter and for short-range forces confined to neighboring sites (long-range forces usually give the Ginzburg-Landau model). It is convenient to define [1] the following four *standard models* in addition to the Ginzburg-Landau model:

(1) The *Ising model* corresponds to $D = 1$ with Hamiltonian $\mathscr{H}$ given by

$$\mathscr{H} = -\sum_{\mathbf{n}}\sum_{\mathbf{i}}{}' J_{\mathbf{i}} S^z_{\mathbf{n}} S^z_{\mathbf{n}+\mathbf{i}} \qquad (5)\ ,$$

where $S^z_{\mathbf{n}}$ is the z-component of spin on the site at $\mathbf{n}$. The exchange parameter J_i couples spins on sites at $\mathbf{n}$ and at $\mathbf{n}+\mathbf{i}$ where $\mathbf{i}$ is a nearest-neighbor vector. The second sum is given a prime to indicate that the sum over $\mathbf{i}$ is to be restricted so that each pair of interacting spins is only included once. The order parameter for the ferromagnetic case (J>0) is $\sum_{\mathbf{n}} S^z_{\mathbf{n}}$, which is one-dimensional. This universality class is appropriate for Ising magnetics, liquid-gas transitions, ordering in alloys, and mixing in liquids.

(2) The *X-Y model* corresponds to $D = 2$ with Hamiltonian X given by

$$\mathscr{H} = -\sum_{\mathbf{n}}\sum_{\mathbf{i}}{}' J_{\mathbf{i}} (S^x_{\mathbf{n}} S^x_{\mathbf{n}+\mathbf{i}} + S^y_{\mathbf{n}} S^y_{\mathbf{n}+\mathbf{i}}) \qquad (6)$$

Here the spins have two components *(x* and y) and the order parameter is the vector sum of these spins, which is two-dimensional. This universality class applied to "easy-plane" magnetics and to superfluidity in liquid ^{4}He.

(3) The *Heisenberg model* corresponds to $D = 3$ with Hamiltonian

$$\mathscr{H} = -\sum_{\mathbf{n}}\sum_{\mathbf{i}}{}' J_{\mathbf{i}}(S^x_{\mathbf{n}}S^x_{\mathbf{n}+\mathbf{i}} + S^y_{\mathbf{n}}S^y_{\mathbf{n}+\mathbf{i}} + S^z_{\mathbf{n}}S^z_{\mathbf{n}+\mathbf{i}}) \tag{7}$$

Here the spin is a three-dimensional vector. The model applies to isotropic magnetic materials.

(4) The *spherical model* corresponds to $D = \infty$. This model assumes the spin $S_{\mathbf{n}}$ has an infinite number of dimensions. It does not seem to correspond to any actual system, but it has the attraction of being solvable exactly.

It is an unfortunate fact that, although a number of exact solutions are available for static critical properties, almost all the critical transitions that occur in the real world correspond to universality classes that cannot be solved exactly. It is useful, however, to list some universality classes within which models have been solved:

1. All cases in one dimension *(d* = 1). Unfortunately these do not show continuous phase transitions and are of limited use for our present purpose.
2. All cases in four or more dimensions *(d* > 4). These give the Ginzburg-Landau solution.
3. The Ising model in two dimensions *(d = 2, D* = 1). This is Onsager's famous solution (Onsager 1944), which gives a continuous phase transition and represents the earliest major progress beyond the Ginzburg-Landau model.
4. The spherical model *(D* = ∞) in any number *d* of dimensions (Berlin and Kac 1952, Stanley 1968).
5. All cases in which the range of the interactions is infinite; these follow the Ginzburg-Landau model.

Many actual critical systems correspond to the three-dimensional case *(d* = 3) and it is an unfortunate fact that the commonly occurring universality classes in three dimensions have not been solved exactly.

2.3 Scaling theory

Although universality has simplified the problem, the direct approach of solving the standard models for commonly occurring universality classes is still too difficult. Below we discuss a less rigorous *scaling approach* using the magnetic language of the standard models. The idea is that the functional form of the free energy should not change if we change the length scale, because it depends on two parameter only: the temperature and the magnetic field.
It can be shown that scaling of the length parameters transforms Gibbs free energy and the correlations length according to the following equations:

$$l^d G(t,h) = G(l^y t, l^x h), \tag{8}$$

$$l^{-1}\xi(t,h) = \xi(l^y t, l^x h) \tag{9}$$

where l is the scaling factor of the length; $t = \frac{T - T_C}{T_C}$ and $h = \frac{g\mu_B H_e}{k_B T}$ are dimensionless temperature and effective magnetic field, respectively.

It can be easily shown that $y = \nu^{-1}$.

The solution of scaling equations (8)-(9) provides the relationships between critical exponent known as *scaling laws*. The importance of scaling theory is that it expresses all the static critical exponents in terms of just two parameters, x and y. The theory gives no information about the values of x and y, so it does not predict critical exponents absolutely.

The scaling laws can be classified effectively into three groups [1]:

1. Scaling laws that put critical exponents above T_c equal to those below T_c:

$$\gamma = \gamma' \quad (10),$$

$$\alpha = \alpha' \quad (11),$$

$$\nu = \nu' \quad (12)$$

2. Scaling laws that do not involve the dimensionality d:

$$\gamma' = \beta(\delta - 1) \quad (13),$$

$$\gamma' + \alpha' + 2\beta = 2 \quad (14),$$

$$(2 - \eta)\nu = \gamma \quad (15)$$

3. Scaling laws that involve the dimensionality d:

$$2 - \alpha = \frac{d}{y} = d\nu \quad (16),$$

$$\beta = (d - x)\nu \quad (17),$$

$$x = (d - x)\delta \quad (18),$$

$$\gamma' = \frac{2x - d}{y} = (2x - d)\nu \quad (19)$$

These are known as the *hyperscaling laws*.

Both the second and third of these categories correspond to replacing exact inequality relationships mentioned in the section 2.1 by equalities. The Ising model in two dimensions obeys all the scaling laws, but the Ginzburg-Landau theory only obeys the hyperscaling laws in four dimensions. For the spherical model, hyperscaling is satisfied if $d \leq 4$ but fails in more than four dimensions.

In Table 2., approximate values of critical exponents are given for various models. For the Ginzburg-Landau, the spherical, and the two-dimensional Ising models the exponents are exact, while for the other models values are taken from approximate calculations of γ and ν and these are used to calculate the other exponents using the scaling laws. Although no scaling law involves the spin-dimensionality, *D,* it is apparent from the table that the critical exponents do change as *D* changes. The scaling parameter y (=ν-1) depends markedly on D, while x depends on weakly, if at all, on *D.* Dashes indicate quantities that do not follow power laws in the critical region.

Table 2. Approximate Values of Critical Exponents for Various Models [1].

Model	Ginzburg-Landau	Ising	Ising	X-Y	X-Y	Heinsenberg	Spherical
D	any	1	1	2	2	3	∞
d	any	2	3	2	3	3	3
γ	1.0	1.75	1.2378 ±0.0006	—	1.316 ±0.009	1.388 ±0.003	1
ν	0.5	1	0.6312 ±0.0003	—	0.669 ±0.007	0.707 ±0.003	1
x	—	1.875	2.481 ±0.001	—	2.484 ±0.009	2.482 ±0.005	2.5
α	—	0	0.106	—	-0.01	-0.121	-1
β	0.5	0.125	0.326	—	0.345	0.367	0.5
δ	3	15	4.78	15	4.81	4.78	5
η	0	0.25	0.039	0.25	0.03	0.037	0

More involved theoretical consideration is based on the so-called renormalization group theory. Similar to the scaling approach, the basic assumption is that changes in length scale merely alter the parameters of the Hamiltonian and not the underlying form of the Hamiltonian. The idea is to apply a continuous family (group) of transformations to the Hamiltonian. These transformations establish correspondences between sets of parameters defining physically different states. Such a correspondence between parameters is referred as

a renormalization of parameters, that specified the name of the approach. A detailed description of this method can be found for instance in the book [6].

2.4 Critical dynamics

As we have already mentioned, there is a slowing down of dynamic processes near critical points, and the response times tend to infinity at the critical point. These phenomena can alternatively be described in terms of certain transport properties becoming infinite at the critical point. It is found experimentally that these divergences follow simple power laws, just as in the case of static properties, so that critical exponents may be defined for dynamic properties. These exponents only follow universality if universality classes are expanded to specify conservation laws as well as the conditions for universality of static exponents. In this section we show that the scaling laws can be extended to include dynamic properties.

It is convenient to discuss the dynamics in terms of the time-dependent correlation function. For this purpose, let us introduce the variable τ for time in order to distinguish time from reduced temperature t. Let us assume that we have spins $\mathbf{S}$ of dimensionality D lying on a lattice with sites $\mathbf{R}$ in d dimensions. Then the spin correlation function $G(\mathbf{R},t,\tau)$ is defined by:

$$G(\mathbf{R},t,\tau) = \left\langle \mathbf{S_0}(0)\cdot \mathbf{S_R}(\tau)\right\rangle - \left\langle \mathbf{S_0}\right\rangle \cdot \left\langle \mathbf{S_R}\right\rangle \qquad (20),$$

where we assume that the average value of $\langle \mathbf{S_R}\rangle$ is time independent, and $\langle \mathbf{S_R}(\tau)\rangle$ is the value of $\mathbf{S}$ on site $\mathbf{R}$ at time τ. For most properties of interest, this function is averaged over all origin sites, $\mathbf{0}$, and where appropriate it will be assumed implicitly that this average is to be taken.

In order to turn to the frequency and momentum variables one can perform the Fourier transformation of Eq.(20):

$$\begin{aligned} G(\mathbf{q},t,\omega) &= \int_{-\infty}^{\infty} \exp(i\omega\tau)\left(\sum_{\mathbf{R}} \left\langle \mathbf{S_0}(0)\cdot \mathbf{S_R}(\tau)\right\rangle - \left\langle \mathbf{S_0}\right\rangle \cdot \left\langle \mathbf{S_R}\right\rangle \right) d\tau = \\ &= \int_{-\infty}^{\infty} \exp(i\omega\tau) G(\mathbf{q},t,\tau) d\tau \end{aligned} \qquad (21)$$

It is usual to split this function into a static part $G(\mathbf{q},t)$ and a dynamic part. So that:

$$G(\mathbf{q},t,\omega) = G(\mathbf{q},t) F(\mathbf{q},t,\omega) \frac{\omega}{1-\exp(-\hbar\omega/k_B T)} \qquad (22)$$

One can show [1], that at the critical point ($t = 0$) the static correlation function obeys to the

scaling law

$$G(\mathbf{q},t=0) \propto |\mathbf{q}|^{\eta-2}, \quad (T=T_c) \tag{23}$$

(cf. expression for the correlation function in real space in Table.1)

Away from the critical temperature, the Fourier transform of the static correlation function leads to complicated analytic forms. In three dimensions, it is a fair approximation to put $\eta = 0$ to obtain the Lorentzian form:

$$G(\mathbf{q},t) \propto \frac{1}{\left(1/\xi\right)^2 + q^2} \tag{24}$$

If one wishes to go beyond this, more precise approximations can be used [1]:

$$G(\mathbf{q},t) \propto \frac{1}{\left(\left(1/\xi\right)^2 + \psi q^2\right)^{1/\psi}}, \text{ with } \psi = \frac{1}{1-\frac{\eta}{2}} \tag{25},$$

or

$$G(\mathbf{q},t) \propto \frac{\left(1/\xi\right)^{\eta}}{\left(1/\xi\right)^2 + q^2} \tag{26}$$

The dynamic part of the correlation function described by Eq.(22) consists of a spectral weight function $F(\mathbf{q},t,\omega)$ and a detailed balance factor $\frac{\omega}{1-\exp(-\hbar\omega/k_B T)}$ (in the critical region, it is approximately constant). The spectral weight function is normalized so that

$$\int_{-\infty}^{\infty} F(\mathbf{q},t,\omega)\,d\omega = 1 \tag{27}$$

In the spirit of scaling theory, we postulate that each volume in reciprocal space of size **q** at temperature *t* in zero magnetic field has a characteristic frequency $\omega_c(\mathbf{q}, t)$. *A* convenient definition of the characteristic frequency is the median frequency, which divides the area under the spectral weight function in half; that is, we define ω_c by:

$$\int_{-\omega_c}^{\omega_c} F(\mathbf{q},t,\omega)\,d\omega = 1/2 \tag{28}$$

Application of the scaling approach to the dynamical correlation function results in the following relation for the characteristic energy at the critical point [1]:

$$\omega_c(\mathbf{q},t=0) \propto q^z, \quad (T=T_c) \tag{29}$$

2.5 Neutron scattering formalism

Since correlation functions determine the scattering properties of a system, critical neutron scattering can be used to measure the correlation function in the critical region. The formalism of the neutron scattering has been discussed in detail in the previous lectures. Here we remind the basic formulae in connection with measurement of the critical scattering following mainly to notation of [1].

The value measured in the neutrons scattering experiment is the partial differential cross section $\frac{d^2\sigma}{d\Omega dE'}$. In case of the system of N nuclei with the same scattering length b:

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{k'}{k} Nb^2 S(\mathbf{Q},\omega) \tag{30}$$

where k and k' are initial and final wave vectors of the neutron, $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$, and $S(\mathbf{Q},\omega)$ is scattering function. $S(\mathbf{Q},\omega)$ can be written in the form:

$$\begin{aligned} S(\mathbf{Q},\omega) = (hN)^{-1} \sum_{jl} \int_{-\infty}^{\infty} \exp(i\omega\tau) d\tau \times \\ \times \left\langle \exp(-i\mathbf{Q}\cdot\mathbf{R}_l(0)) \exp(-i\mathbf{Q}\cdot\mathbf{R}_j(\tau)) \right\rangle \end{aligned} \tag{31},$$

so that it is proportional to the Fourier transform of a correlation function between the position vector of the l^{th} atom at time zero and the j^{th} atom at time τ.

Similarly, for magnetic scattering

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{k'}{k} \frac{N}{\hbar} (\gamma r_0)^2 \left| f(\mathbf{Q}) \right|^2 \sum_{\alpha\beta} (\delta_{\alpha\beta} - \hat{Q}_\alpha \hat{Q}_\beta) S^{\alpha\beta}(\mathbf{Q},\omega) \tag{32},$$

where γ is the gyromagnetic ratio for neutron, r_0 is the classical radius of the electron, $f(\mathbf{Q})$ is the magnetic form factor, $\hat{Q}_\alpha$ is the direction cosine of $\mathbf{Q}$ along the α axis.

The scattering function

$$S^{\alpha\beta}(\mathbf{Q},\omega) = (2\pi)^{-1}\sum_{\mathbf{R}}\exp(i\mathbf{Q}\cdot\mathbf{R})\int_{-\infty}^{\infty}\exp(i\omega\tau)\left\langle S_{\mathbf{0}\alpha}(0)S_{\mathbf{R}\beta}(\tau)\right\rangle d\tau \qquad (33),$$

represents the Fourier transform of a dynamic correlation function.

If one needs to measure the static correlation function, it is necessary to integrate $S^{\alpha\beta}(\mathbf{Q},\omega)$ over the frequency in the infinite range:

$$G^{\alpha\beta}(\mathbf{Q},t,h) = \int_{-\infty}^{\infty} S^{\alpha\beta}(\mathbf{Q},t,\omega)d\omega \qquad (34),$$

There is a problem with this process, however, in that the measurements cannot, in practice, be taken over an infinite range of ω, so that we cannot evaluate the contribution to the integral for all values of ω. Accurate measurements can only be taken if there is a good signal-to-noise ratio and if the weight of the integral is concentrated over a relatively narrow range of ω. Fortunately, this latter condition often might be reasonably satisfied near the critical point, since the characteristic frequency ω_c tends to zero at the critical point.

There is a second way in which we can attempt to extract the static correlation function from neutron-scattering measurements. This involves measuring the differential magnetic cross section rather than the partial differential magnetic cross section. That is, we measure all the neutrons scattered into solid angle *dΩ,* without regard to energy. This is easy to do in practice since it just involves using a detector with angular size *dΩ.* What we will observe will be the integral of the right-hand side of Eq. (32) over the scattered neutron energy *dE′* at fixed scattering angle:

$$\frac{d\sigma}{d\Omega} = \frac{N(\gamma r_0)^2}{k\hbar}\int_0^{\infty} dE'\,k'\left|f(\mathbf{Q})\right|^2\sum_{\alpha\beta}(\delta_{\alpha\beta} - \hat{k}_\alpha\hat{k}_\beta)S^{\alpha\beta}(\mathbf{Q},\omega) \qquad (35),$$

If all the weight of the scattering function $S^{\alpha\beta}(\mathbf{Q},\omega)$ is at low frequencies ($\hbar\omega << E$), it will be a good approximation to evaluate the integral with *k′* and **Q** held constant (such that |**k'**| = |**k**|). This then gives

$$\frac{d\sigma}{d\Omega} = \frac{N}{\hbar}(\gamma r_0)^2\left|f(\mathbf{Q})\right|^2\sum_{\alpha\beta}(\delta_{\alpha\beta} - \hat{Q}_\alpha\hat{Q}_\beta)G^{\alpha\beta}(\mathbf{Q},t,h) \qquad (36),$$

which enables the spin correlation function *G(**Q**,t,h)* to be measured. Similar formula can be obtained for the nuclear scattering.

The approximation that was made in putting $\hbar\omega << E$ for all ω that contributes significantly to the scattering in the correlation function is known as the *static* (or *quasistatic*) approximation, so called because it gives the "static" correlation function. It can be shown that the approximation is equivalent to assuming that the spin vectors **S** have not had the time to change during the time that it takes the neutron to cross an atom, so that the diffraction pattern corresponds to a "static" set of spins. The advantage of measuring spin correlation functions in this way is that the differential cross section can be measured more easily and more accurately than the partial differential cross section. The disadvantage is that we do not normally know how good the static approximation is.

It is useful also to note that one has to distinguish between the measurements using static approximation and the measurements with fixed energy transfer E=0 (elastic scattering). The integrated cross section $\frac{d\sigma}{d\Omega}$ contains both delta-function Bragg scattering from the long-range order at magnetic reciprocal lattice points and *diffuse* scattering from the short-range order. The elastic scattering $\frac{d^2\sigma}{d\Omega dE'}(\omega = 0)$ consists of only the Bragg scattering part of the differential scattering. From the formal point of view, this means that in the static approximation, the differential scattering gives spin correlations at the time τ= 0, while the elastic scattering gives correlation over infinite times.

3 Measurement of critical scattering

3.1 Classical magnetic materials

In this section we turn to the experimental examples of the critical scattering, and we start from the classical magnetic materials. In Fig. 4:(a,b) the critical scattering patterns for one-dimensional Ising system K_2CoF_4 just above and below T_c [7] is shown. The data have been obtained in the quasistatic approximation discussed in the previous section.

Near the critical point the correlation length becomes large and the correlation function is sharply peaked in reciprocal space. In this case the effects of experimental resolution may become marked and this must be taken into account in interpreting the data. The determination of the correlation length ξ and its critical exponent ν from the scattering depends on knowledge of the form of the correlations, and normally this form is not known, so an approximate form must be assumed. The experimental data above T_c (Fig. 4:(a)) have been fitted by the Lorentzian peak convoluted with the resolution function, and the quality of fitting (shown by solid line) is excellent. Fitting by Lorentzian below T_c (Fig. 4:(b)) seems to be not too bad, too, but the better agreement with experiement can be achieved using Tarko-Fisher peak shape (dashed line in Fig. 4:(b)). This results in the different values of the critical exponent for the correlation length, those are $\nu' = 0.86 \pm 0.06$ with χ^2=1.10 and $\nu' = 1.12 \pm 0.13$ with χ^2=0.40, for the Lorentzian and Tarko-Fisher shapes, respectively (cf. the value above T_c $\nu = 1.02 \pm 0.05$ that is expected to be equal to $\nu = 1.12 \pm 0.13$).

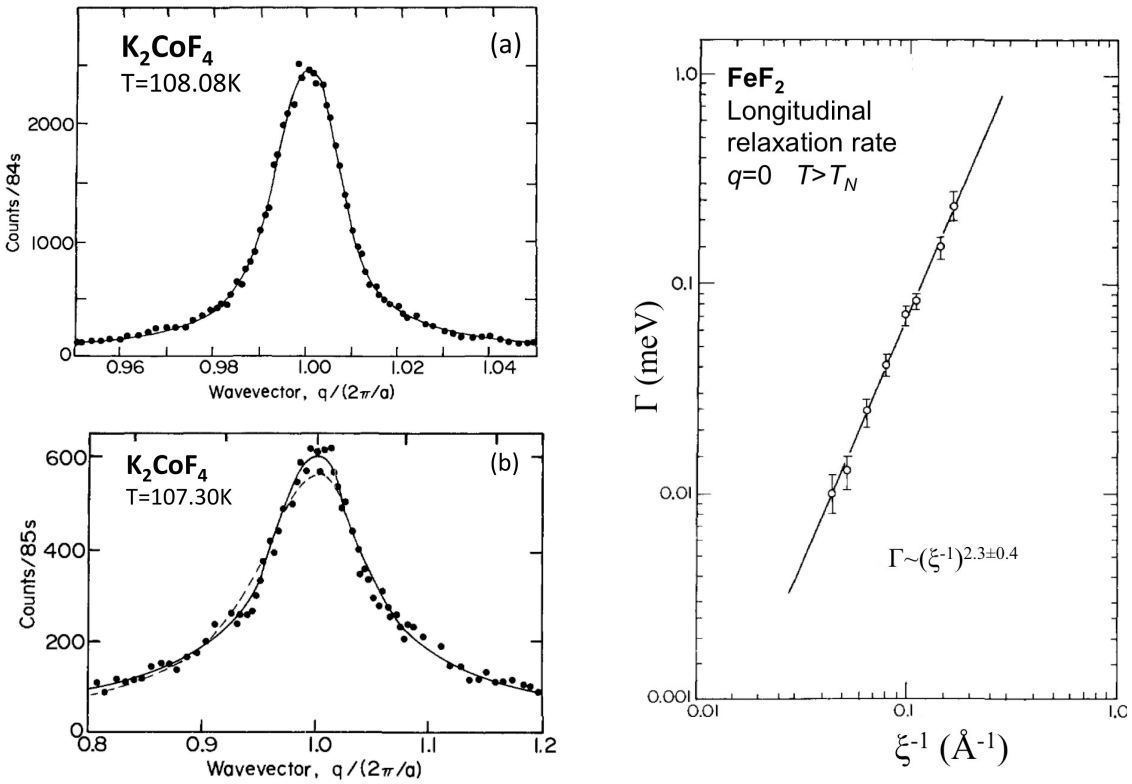


Fig. 4: *(a)-(b): The critical scattering observed in* ***K_2CoF_4*** *at two temperatures above and below T_c (T_c = 107.72K) for* ***Q*** *= (q, 0, 0.45) (adapted from [7]).*
(c): Logarithmic plot of the characteristic frequency for longitudinal fluctuations versus the inverse correlation length for ***FeF_2*** *at temperatures above T_N and at q = 0 (adapted from [8]). The solid line represents the best fit to the data of a power law with critical exponent z=2.3 ± 0.4.*

The determination of the dynamical exponent is illustrated in Fig. 4:(c) for three-dimensional Ising magentics FeF_2. The longitudinal dynamic correlation in this system show a peak of Lorentz shape at ω = 0 [8]: $F_{//}(\omega) = \dfrac{\Gamma_{//}}{\Gamma_{//}^2 + \omega^2}$, so that the characteristic energy $\Gamma_{//}(\mathbf{q},t)$ follows to the scaling dependence $\Gamma_{//} \propto \mathsf{k}^z$, where $\mathsf{k} = \xi^{-1}$ is inverse correlation length. The value of z obtained from the linear fitting of the data is z = 2.3±0.4.

3.2 Magnetic systems with quantum phase transitions

More topical object for the magnetic critical scattering is quantum phase transitions. In contrast to a classical phase transition at nonzero temperatures, driven by temperature as a control parameter with thermal fluctuations, a quantum phase transition is driven by a control parameter other than temperature, e.g., external pressure, magnetic field or doping, at absolute

zero, with quantum-mechanical fluctuations. Such a control parameter tunes a system at the zero-temperature boundary from an ordered ground state towards a non- ordered state crossing a quantum critical point. Although this definition of a quantum phase transition is strictly valid only for T=0, sufficiently close in temperature to this critical point the system's behavior is still determined by the quantum critical point. The nature of a classical transition at some finite temperature T_c is characterized by a diverging correlation length and correlation time approaching the transition. Such fluctuations of the order parameter are associated with a frequency ω* that vanishes at the transition. A quantum system at finite temperatures behaves in a classical way, if the temperature exceeds the fluctuation frequencies. This argument also shows that quantum phase transitions, where $T_c = 0$, are qualitatively different, and their critical fluctuations must be treated quantum mechanically.

In the methodology of the path-integral formulation of quantum mechanics, a *d*-dimensional quantum system (T=0) can be viewed as a (d+z)-dimensional classical system, where z is the scaling exponent for the dynamics. This allows the application of general ideas of finite-temperature critical points to quantum critical phenomena. In particular, this means that in some cases one can apply the models for dimension $d > 4$ those were mentioned to be solvable exactly. Since in experiment only the behavior for $T > 0$ can be studied, the identification of the underlying quantum critical point in the phase diagram relies on the findings of scaling behavior, e.g., for temperature or frequency. Quantum phase transitions are closely connected with another interesting effect – non-Fermi-liquid behavior (for the relatively recent review on this problem see [9]).

An example of the critical scattering in the system with magnetic quantum phase transition is shown in Fig. 5: for $CePd_{1-x}Rh_x$ [10]. In this compound the control parameter driving the phase transition is doping (i.e. concentration of Rh, x). The critical composition corresponds to x=0.85. The magnetic component of the scattering function at temperatures between 5K and 254K (Fig. 5:)) demonstrate a well-pronounced quasielastic signal, which is nearly independent on the neutron energy loss side and follows the thermal population factor on the neutron energy gain side.

Fig. 5:) shows the scaling of the imaginary part of the dynamic magnetic susceptibility $\chi''(\mathbf{Q},\omega,T)$ obtained using standard relation between $\chi''(\mathbf{Q},\omega,T)$ and the scattering function $S(\mathbf{Q},\omega,T)$:

$$\chi''(\mathbf{Q},\omega,T) = 2\pi\left(\frac{\mu_B}{\gamma r_0}\right)^2 S(\mathbf{Q},\omega,T) \tag{37}$$

It was found that $\chi''(\mathbf{Q},\omega,T)$ can be described by the scaling relation

$$\chi''(\omega,T)T^{\alpha} = \left(\frac{T}{\omega}\right)^{\alpha} tahh\left(\omega/T\beta\right) \tag{38},$$

suggested initially for $UCu_{5-x}Pd_x$ system [11].

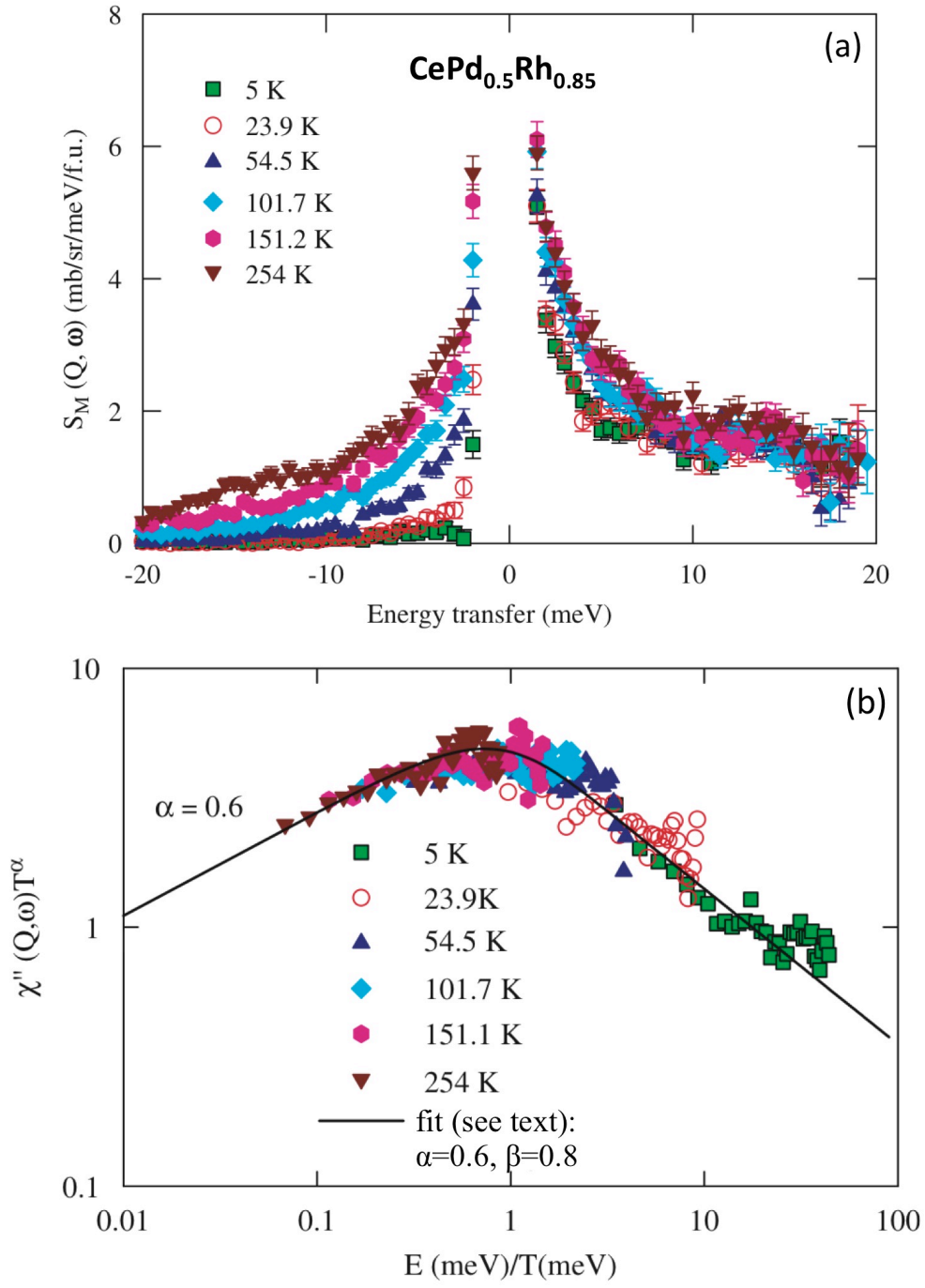


Fig. 5: *The critical scattering and scaling in $CePd_{0.15}Rh_{0.85}$ (adapted from [10]). (a) Inelastic magnetic neutron spectra at various temperatures. (b) E/T scaling plot of the magnetic scattering. The solid line is a fit of the experimental data by the scaling function (see text).*

The solid line in Fig. 5:) represents the fit of the experimental data according to Eq.(38). The critical exponents obtained from the fitting are: α=0.6, β=0.8.

We have also to mention another specific feature of the quantum phase transitions. As the control parameter is other than temperature, one can trace the exponent behaviour of some physical properties with a base argument like pressure or magnetic field. For instance, the divergence of the correlation length near the phase transition in $Pr2CuO_4$ is described [12] by the power law

$$\xi \sim \left(1 - \frac{H}{H_C}\right)^{-\nu} \quad \text{for } H<H_C \tag{39},$$

and

$$\xi \sim u\left(\frac{H}{H_C} - 1\right)^{-\nu} \quad \text{for } H>H_C \tag{40},$$

where H_C is the critical field, and u is a factor to describe the asymmetry of the critical scattering with respect to H_C.

3.3 Phase transitions driven by the soft modes

In conclusion, we turn to another extended field where the critical scattering has been efficiently used. This is the displacive structure phase transitions driven by the soft phonon mode. Initially, the soft-mode concept was developed to describe the origin of ferroelectricity, but then it has since been successfully applied to explain a wide variety of structural transformations. Inelastic neutron scattering has been successfully employed to characterize the soft-mode dynamics in rock-salt semiconductors (PbTe, SnTe etc.) and α-quartz structures, as well as for a number of ferroelectrics such as $BaTiO_3$, $KTaO_3$, $PbTiO_3$, $KNbO_3$. Detailed neutron studies have been performed for and many of the perovskites, such as $SrTiO_3$, $KMnF_3$, $CsPbCl_3$, and $LaAlO_3$ (for a basic review see, e.g., [13]).

In essence, soft mode theory states that the eigenvector of a certain lattice vibrational mode describes the pattern of atomic displacements necessary to transform the solid from its high-temperature phase into the low-temperature phase. As temperature is reduced towards the transition temperature, T_c, the frequency of this mode decreases, eventually vanishing at T_c and resulting in the structural transformation. The transition can be viewed as a “freezing” of the relevant vibrational mode. In ferroelectrics the soft mode occurs at $q \sim 0$. In the case of antiferroelectrics, the soft mode is located at the zone boundary. When a mode freezes out at a zone-boundary point, a new Bragg peak generally appears at that position in reciprocal space. The intensity of this superlattice peak is proportional to the square of the order parameter and, consequently, it increases as the temperature is reduced. A study of the intensity of these new peaks is important in establishing the low-temperature structure.

For more detailed analysis of the soft-mode-driven phase transitions we will limit ourselves by considering the results for α-β phase transformation in quartz which occurs at T_0 = 846 K

[14]. In the high temperature β-phase the soft mode is overdamped and located at the Brillouin zone center. Fig. 6:(a) shows inelastic neutron scattering spectra at such a position, **Q** = (1, 0, 3).

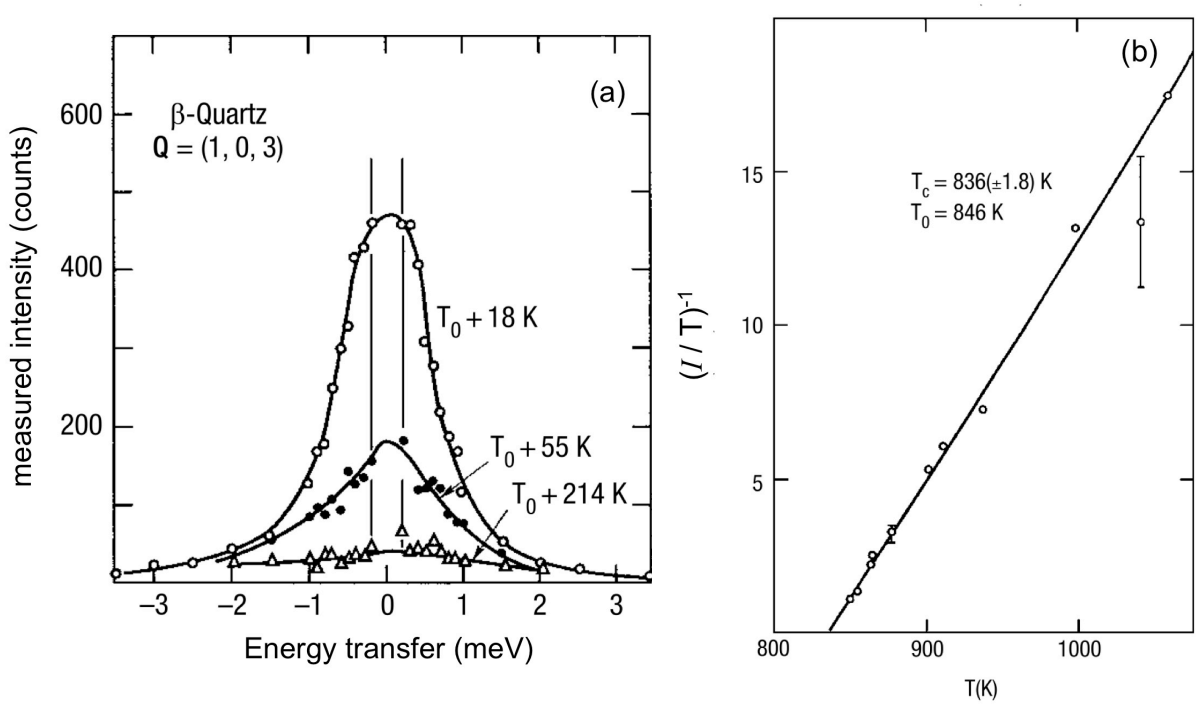


Fig. 6: *(a)-(b): Measurements of the overdamped soft mode at the β → α transition in quartz (adapted from [14]).(a) Inelastic neutron scattering spectra measured at constant **Q** (**Q**=(1,0,3)) for several temperatures above the transition. (b) Inverse of the energy-integrated intensity divided by temperature, plotted versus temperature. The straight line is a fit (see text).*

The scattering from the soft mode appears as a broad Lorentzian, easily separated from the sharp Bragg peak, with an intensity that grows rapidly as T_0 is approached.
The partial differential cross section in this case can be written [14] as:

$$\frac{d^2\sigma}{d\Omega dE} \propto \left|\frac{k'}{k}\right|^2 \left|F_{in}(\mathbf{Q})\right|^2 \left[1+n(\omega)\right] A(\omega) \qquad (41),$$

where $F_{in}(\mathbf{Q})$ is the inelastic structure factor for the normal phonon mode, $n(\omega)$ is the phonon occupation number.

The spectral correlation function

$$A(\omega) = \frac{1}{\pi} \frac{2\omega\Gamma}{(\omega_0^2 - \omega^2)^2 + (\omega\Gamma)^2} \tag{42}$$

where ω_0 is the energy of the soft mode. Taking into account strong damping of the soft mode ($\Gamma^2 > 2\omega_0^2$), one can derive that the experimental intensity integrated over ω is:

$$I = \int \frac{d^2\sigma}{d\Omega dE} dE \propto \left|F_{in}(\mathbf{Q})\right|^2 \frac{T}{\omega_0^2} \tag{43}$$

From this equation, one see that the T/I ratio should be proportional to ω_0^2; this quantity is plotted versus temperature in Fig. 6:(b). The straight line through the data points is a fit to the equation

$$\omega_0^2 = a(T - T_c) \tag{44}$$

The fit yields $T_c = 836 \pm 1.8$ K, that is 10 K less than the actual first-order transition temperature, T_0.

References

[1] M. F. Collins, Magnetic Critical Scattering (Oxford University Press, New York, 1989)

[2] http://en.wikipedia.org/wiki/Triple_point

[3] J. E. Noakes, N. E. Tornberg, and A. Arrott, J. Appl. Phys. 37, 1264 (1966)

[4] H. E. Stanley, Introduction to Phase Transitions and Critical Phenomena (Oxford University Press, London, 1971)

[5] G. S. Rushbrooke, J. Chem. Phys. 39**,** 842 (1963)

[6] P. Pfeuty, G. Toulouse, Introduction to the Renormalization Group and to Critical Phenomena (Wiley, New York, 1977)

[7] R. A. Cowley, M. Hagen, and D. P. Belanger, J. Phys. C.: Solid State Phys., 17, 3763 (1984)

[8]M. T. Hutchungs, M. P. Schulhof, H. J. Guggenheim, Phys. Rev. B 5, 154 (1972)

[9] G. R. Stewart, Rev. Mod. Phys. 73 797 (2001)

[10] D. T. Adrroja, J.-G. Park, Kwang-Hyun Jang, H. C. Walker, K. A. McEven, T. Takabatake, J. Magn. Magn. Mater. 310, 858 (2007)

[11] M.C. Aronson, R. Osborn, R. A. Robinson, J. W. Lynn, R. Chau, C. L. Seaman, M. B. Maple, Phys. Rev. Lett 75, 725 (1995)

[12] A. Ivanov, D. Petitgrand, J. Magn. Magn. Mater. 272-276, 220 (2004)

[13] J. F. Scott, Rev. Mod. Phys. 46 83 (1974)

[14] J. D. Axe, G. Shirane, Phys. Rev. B 1, 342 (1970)

D 7 Quasielastic Scattering [1]

Joachim Wuttke
Jülich Centre for Neutron Science at FRM II
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

The term 'quasielastic scattering' designates a limiting case of inelastic scattering that is close to elastic scattering. There are different ways to specify what 'close' means. In a wider sense, it is only required that the energy transfer be small compared to the incident energy of the scattered particles. This is how the word 'quasielastic' originally was coined in nuclear physics. In a narrower sense, a *quasielastic line* is a spectral feature centered around the elastic peak. A very special use or abuse of the term 'quasielastic' shall only be mentioned in passing: *quasielastic light scattering*, also called *dynamic light scattering* or (much more to the point) *photon correlation spectroscopy*.[2]

Quasielastic neutron scattering (QENS) is used to study atomic and molecular modes of motion that are slower than the phonons typically studied on a triple-axis spectrometer. A lower bound for the energy transfers is set by the resolution of available spectrometers. In the following, we will only discuss energy-resolved measurements on time-of-flight or backscattering spectrometers, refering to another lecture (D8) for the spin-echo method, which is special in several respects (implicit Fourier transform, preference for coherent scattering, focus on small q). Accordingly, we are concerned with energy transfers of the order of μeV to meV, corresponding to time scales of ns to ps. Motion on this time-scale involves molecular reorientations and certain slow oscillations as well as jump processes that add up to diffusion or relaxation.

Depending on the isotopic composition of the sample, neutron scattering can be dominated by coherent or incoherent contributions (lecture A4). Coherent scattering conveys richer information, but is more difficult to interpret. On time-of-flight and backscattering spectrometers, by far the majority of experiments is concerned with incoherent scattering. In particular, whenever a sample contains hydrogen, this element is likely to dominate the total cross section, except if the material is fully deuterated. In this lecture we will concentrate on the most simple and most important use of QENS: *incoherent scattering* by *hydrogen*.

2 Basics

2.1 Self correlation functions

Incoherent neutron scattering measures *self correlations* of *tagged particles*. Assuming that there is just one noteworthy species of scatterers, the double differential cross section can be written as[3]

$$\frac{\partial^2\sigma}{\partial\Omega\partial\omega} = \frac{\sigma_{\mathrm{inc}}}{4\pi}\frac{k_{\mathrm{f}}}{k_{\mathrm{i}}}S_{\mathrm{inc}}(\mathbf{Q},\omega). \tag{1}$$

For the remainder of this lecture, the subscript 'inc' will be omitted. Let us summarize a few relations from lecture A5. The *scattering function* $S(\mathbf{Q},\omega)$ is best expressed via a Fourier transform in time,

$$S(\mathbf{Q},\omega) = \frac{1}{2\pi}\int \mathrm{d}t\, \mathrm{e}^{-i\omega t}\, I(\mathbf{Q},t), \tag{2}$$

[2] In photon-correlation spectroscopy, scattered photons are counted regardless of their energy; photon counts are then correlated by some real-time electronics. This has almost nothing in common with quasielastic neutron scattering. The light scattering analogue of quasielastic neutron scattering is called high-resolution inelastic light scattering or *Rayleigh-Brillouin scattering*; to analyse the energy of scattered photons either a Fabry-Perot interferometer or a high-resolution grating spectrometer is used.

[3] Only here, Ω stands for a solid angle. Later on, we will reuse the symbol Ω to designate characteristic frequencies of our model systems.

with the *intermediate scattering function*

$$I(\mathbf{Q},t) = \frac{1}{N}\sum_j \left\langle \mathrm{e}^{-i\mathbf{Q}\hat{\mathbf{r}}_j(0)}\mathrm{e}^{i\mathbf{Q}\hat{\mathbf{r}}_j(t)}\right\rangle . \tag{3}$$

Its initial value is $I(\mathbf{Q},0) = 1$. In the ballistic short-time regime (for times that are short compared to a typical phonon period), it has the expansion

$$I(\mathbf{Q},t) = \frac{1}{N}\sum_j \left\langle \mathrm{e}^{i\mathbf{Q}\mathbf{v}_j(0)}\right\rangle \doteq 1 - \frac{1}{2}\Omega_Q^2 t^2 + \ldots \tag{4}$$

with a frequency Ω_Q that depends on Q and on the mean squared thermal velocity:

$$\Omega_Q^2 := \frac{1}{N}\sum_j \left\langle (\mathbf{Q}\mathbf{v}_j(0))^2\right\rangle = \frac{Q^2}{3}\frac{1}{N}\sum_j \left\langle \mathbf{v}_j(0)^2\right\rangle = Q^2 k_\mathrm{B}T\frac{1}{N}\sum_j \frac{1}{m_j}. \tag{5}$$

Since Ω_Q lies far outside the dynamic range of the spectrometers used for QENS, the short-time expansion has only theoretical importance.

An additional Fourier transform in space,

$$I(\mathbf{Q},t) = \int \mathrm{d}^3 r\, \mathrm{e}^{i\mathbf{Q}\mathbf{r}}\, G_\mathrm{s}(\mathbf{r},t) \tag{6}$$

relates the intermediate scattering function to the self-correlation function

$$G_\mathrm{s}(\mathbf{r},t) = \frac{1}{N}\sum_j \int \mathrm{d}^3 r'\ \left\langle\, \delta(\mathbf{r}-\mathbf{r}'+\hat{\mathbf{r}}_j(0))\, \delta(\mathbf{r}'-\hat{\mathbf{r}}_j(t))\,\right\rangle . \tag{7}$$

Before taking the classical limit, it is advisable to correct for detailed balance, introducing the symmetrized scattering law

$$\tilde{S}(\mathbf{Q},\omega) := \mathrm{e}^{\hbar\omega/2k_\mathrm{B}T} S(\mathbf{Q},\omega). \tag{8}$$

In the classical limit, the position operators commute, and the symmetrized self-correlation function $\tilde{G}_\mathrm{s}$ is approximated by

$$G_\mathrm{s}^\mathrm{cl}(\mathbf{r},t) = \frac{1}{N}\sum_j \left\langle\, \delta(\mathbf{r}-\hat{\mathbf{r}}_j(t)+\hat{\mathbf{r}}_j(0))\,\right\rangle . \tag{9}$$

2.2 Harmonic vibrations and the mean squared displacement

Before turning to the slow, anharmonic modes of motion that are typically studied by QENS, we briefly review scattering by harmonic vibrations (compare lecture B4). This analytically tractable reference case helps us to introduce important concepts like elastic incoherent structure factor and the mean squared displacement, which later on will be generalized for non-harmonic systems.

Vibrations are described in terms of displacements $\mathbf{u}_j(t)$ from equilibrium positions $\mathbf{R}_j$,

$$\mathbf{r}_j(t) = \mathbf{R}_j + \mathbf{u}_j(t). \tag{10}$$

Eq. (3) keeps its form under this transformation, except that the $\mathbf{r}_j$ are substituted by $\mathbf{u}_j$. If the particles j are only subject to harmonic forces, then the Bloch theorem [1] allows the simplification

$$I(\mathbf{Q},t) = \frac{1}{N}\sum_j \mathrm{e}^{-2W_j(\mathbf{Q},\mathbf{0})}\mathrm{e}^{2W_j(\mathbf{Q},t)} \tag{11}$$

with

$$2W_j(\mathbf{Q},t) := \langle\,(\mathbf{Q}\mathbf{u}_j(0))(\mathbf{Q}\mathbf{u}_j(t))\,\rangle\,. \tag{12}$$

In isotropic systems, an orientational average gives

$$2W_j(Q,t) = \frac{Q^2}{3}\,\langle\,\mathbf{u}_j(0)\mathbf{u}_j(t)\,\rangle\,. \tag{13}$$

Each particle partakes in a huge number of oscillatory modes, which quickly run out of phase. Therefore, within little more than one typical phonon period, $2W_j(\mathbf{Q},t)$ approaches 0. In consequence, the intermediate scattering function converges towards the long-time limit

$$I(\mathbf{Q},t\to\infty) = \frac{1}{N}\sum_j \mathrm{e}^{-2W_j(\mathbf{Q},\mathbf{0})} =: f_{\mathbf{Q}}. \tag{14}$$

This value is called the *Debye-Waller factor*, or, if one wants to emphasize the distinction between incoherent and coherent scattering, the *Lamb-Mössbauer factor* or the *elastic incoherent structure factor* (EISF). The Fourier transform of (11) gives

$$S(\mathbf{Q},\omega) = f_{\mathbf{Q}}\delta(\omega) + (1-f_{\mathbf{Q}})S_{\mathrm{inelast}}(\mathbf{Q},\omega). \tag{15}$$

Here the EISF reappears as the amplitude of the elastic delta line, and its complement as the intensity of the inelastic phonon spectrum. In a first approximation, this spectrum is proportional to the *vibrational density of states* (VDOS) $g(\omega)$, divided by ω^2. As soon as one goes beyond the first approximation, the extraction of a VDOS from inelastic neutron scattering data becomes a tour de force of non-linear, interdependent corrections.

But even the elastic intensity contains *some* information about the VDOS: Assuming that there is just one type of scatterers, Eqs. (13) and (14) provide a simple expression for the *mean squared displacement* (MSD)

$$\langle u_x^2\rangle = \frac{\langle\mathbf{u}^2\rangle}{3} = \frac{-\ln f_Q}{Q^2}. \tag{16}$$

On backscattering instruments, one can rather comfortably measure f_Q as function of temperature T (elastic temperature scan with no Doppler modulation of the incident neutron energy). From the slope of $-\ln f_Q(T)$ versus Q^2 one obtains the MSD.

The MSD in turn is a temperature weighed moment of the VDOS,

$$\langle u_x^2\rangle = \frac{\hbar}{6m}\int \mathrm{d}\omega\,\frac{g(\omega)}{\omega}\coth\frac{\hbar\omega}{2k_\mathrm{B}T}. \tag{17}$$

Approximating the VDOS by the *Debye model*,

$$g_\mathrm{D}(\omega) = \begin{cases} \dfrac{9\omega^2}{\omega_\mathrm{D}^3} & \text{for } 0\le\omega\le\omega_\mathrm{D}, \\ 0 & \text{else,} \end{cases} \tag{18}$$

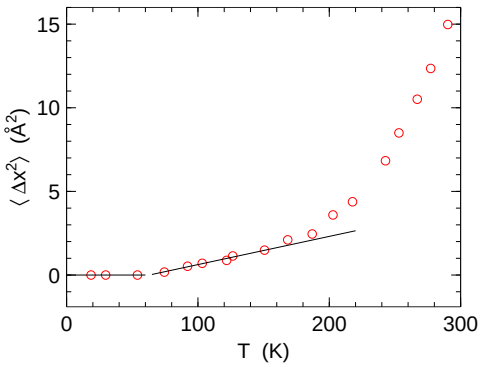


Fig. 1: *Temperature dependence of the mean squared displacement in hydrated myoglobin, measured by thermal neutron backscattering (IN13 at ILL), redrawn from the seminal work by Doster, Cusack, and Petry (1989) [2]. The solid lines indicate the asymptotes (19) expected for harmonic vibrations, with a clear crossover from quantum zero-point oscillations to a linear regime. Anharmonic contribution appear above 150 K, and increase strongly above 180 K. In the last 20 years, this "protein dynamic transition" has been intensely studied, but its nature is still disputed [3].*

and expressing the cut-off frequency $\omega_{\rm D}$ through the equivalent *Debye temperature* $\Theta_{\rm D} := \hbar\omega_{\rm D}/k_{\rm B}$, we find the asymptotes

$$\left\langle u_x^2 \right\rangle \doteq \begin{cases} \dfrac{3\hbar^2}{4mk_{\rm B}\Theta_{\rm D}} & \text{for } T \ll \Theta_{\rm D}, \\ \dfrac{3\hbar^2}{mk_{\rm B}\Theta_{\rm D}}\dfrac{T}{\Theta_{\rm D}} & \text{for } T \gg \Theta_{\rm D}. \end{cases} \tag{19}$$

At $T \simeq \Theta_{\rm D}/4$, the MSD crosses over from temperature-independent quantum zero-point oscillations to a linear temperature dependence. This crossover is indeed regularly observed in the low-temperature part of elastic scans. At higher temperatures, deviations from the linear behavior $\langle u_x^2\rangle \propto T$ reveal the onset of anharmonic motion (Fig. 1).

2.3 Quasielastic broadening; localized versus diffusive motion

The quantum theory of harmonic vibrations has provided us with a clear decomposition (15) of the scattering function into an elastic peak and a broad phonon spectrum. This is no longer the full story if there are non-vibrational degrees of freedom, due for instance to molecular rotation or atomic diffusion. In such cases, the elastic delta function is expected to *broaden* into a quasielastic spectrum. However, if the non-vibrational motion is much slower than typical vibrations, the overall structure of Eq. (15) remains a good approximation:

$$S(\mathbf{Q},\omega) \simeq f_{\mathbf{Q}} S_{\rm QENS}(\mathbf{Q},\omega) + (1-f_{\mathbf{Q}}) S_{\rm inelast}(\mathbf{Q},\omega). \tag{20}$$

If there is no clear separation of time scales, it may be more appropriate to replace $S_{\rm inelast}$ by a convolution $S_{\rm QENS} \otimes S_{\rm inelast}$.

The replacement of $\delta(\omega)$ by $S_{\rm QENS}(\mathbf{Q},\omega)$ does not necessarily mean that there is no more central delta line. In the next section, we will discuss jump and rotator models for which the quasielastic scattering has the form

$$S_{\rm QENS}(\mathbf{Q},\omega) = a_{\mathbf{Q}}\delta(\omega) + b(\mathbf{Q},\omega). \tag{21}$$

The presence of an elastic line ensures that the intermediate scattering function has a nonzero long-time limit $I(\mathbf{Q}, t \to \infty) = f_\mathbf{Q} a_\mathbf{Q}$. This is the defining characteristic of *localized motion*: the scatterers remain confined to a finite region in space. In the opposite case of *long-ranged diffusion*, $I(\mathbf{Q}, t)$ has the long-time limit 0, and S_{QENS} has no delta component.

We also need to generalize the mean squared displacement. It becomes a time-dependent function that can be computed as the second moment of the self-correlation function,

$$\left\langle r^2(t) \right\rangle := \int \mathrm{d}^3 r \, r^2 \, G_\mathrm{s}(\mathbf{r}, t). \tag{22}$$

2.4 Important fit functions

When it comes to fitting QENS data, one usually starts with very simple model functions like a Lorentzian, and in many cases this is all one needs. Let us therefore collect some basic facts about a few fit functions.

Computationally the simplest fit function is the Gaussian. Unfortunately, it has few applications beyond the textbook case of an ideal gas (lecture A5). A normalized Gaussian, as one would use to fit a scattering function, has the standard form

$$\mathcal{G}(\omega; \Gamma) := \frac{1}{\sqrt{2\pi}\Gamma} \mathrm{e}^{-\omega^2/2\Gamma^2}. \tag{23}$$

Its Fourier transform, corresponding to the intermediate scattering function, is also a Gaussian:

$$\tilde{\mathcal{G}}(t; \Gamma) = \mathrm{e}^{-\Gamma^2 t^2/2}, \tag{24}$$

which is perfectly consistent with the short-time expansion (4). The width of a Gaussian spectrum is often expressed by its *full width at half maximum* (FWHM), which can be easily computed as $\sqrt{8 \ln 2}\Gamma$. Intermediate scattering functions are usually characterised by a *mean relaxation time*,

$$\langle \tau \rangle := \int_0^\infty \mathrm{d}t \, I(Q, t) = \pi S(Q, 0). \tag{25}$$

For the Gaussian, one finds $\langle \tau \rangle = \sqrt{\pi/2}/\Gamma$.

Diffusion, jump processes, and rotations can all be modelled by variants of the *master equation*, leading to a *Lorentzian* scattering function. The normalized Lorentzian, in mathematics known as the *Cauchy distribution*, has the form

$$\mathcal{L}(\omega; \Gamma) := \frac{1}{\pi} \frac{\Gamma}{\Gamma^2 + \omega^2}. \tag{26}$$

Its Fourier transform is a simple exponential,

$$\tilde{\mathcal{L}}(t; \Gamma) = \mathrm{e}^{-\Gamma t}. \tag{27}$$

The FWHM is 2Γ, and the mean relaxation time $\langle \tau \rangle = 1/\Gamma$. The exponential (27) is *not* compatible with the short-time expansion of $I(Q, t)$: the master equation does not adequately capture the ballistic short-time regime. The limited range of validity of Lorentzian fits is also obvious from the fact that the Cauchy-Lorentz distribution does not possess a second moment, whereas a sum rule relates the second moment of $S(Q, \omega)$ to the velocity autocorrelation function.

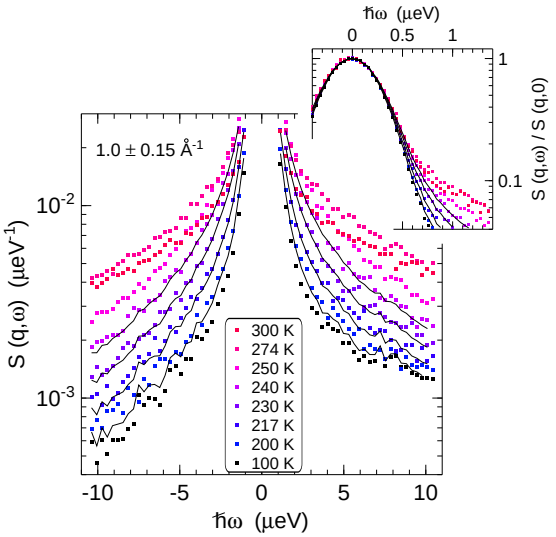


Fig. 2: *Spectra of the hydrated protein c-phycocyanin, measured on the backscattering spectrometer SPHERES of JCNS. Solid lines are fit with a Kohlrausch-Williams-Watts function ($\beta = 0.5$), numerically convolved with the resolution measured at 100 K (black symbols); abover 240 K, this function becomes inadequate, and measurements over a wider energy range are needed to establish a physically valid description of $S(Q,\omega)$. As in many other QENS experiments, quasielastic scattering first appears deep in the wings of the resolution function, whereas no broadening can be seen at half maximum (inset). Redrawn from [4].*

Another important fit function is the *stretched exponential function*, which is defined in the time domain as

$$\tilde{\mathcal{K}}_\beta(t;\tau) := \mathrm{e}^{-(t/\tau)^\beta}. \tag{28}$$

Its Fourier transform, often called the *Kohlrausch-Williams-Watts* function, must be computed numerically except in a few special cases ($\beta = 0.5, 1, 2$). With typical values of β between 0.4 and 0.8, it is often used to describe relaxation in viscous liquids; with the special value $\beta = 0.5$, it includes a key result of the Rouse model for polymer motion. Its characteristic relaxation time is $\langle\tau\rangle = \tau\Gamma(1/\beta)/\beta$ where $\Gamma()$ is the gamma function.

2.5 Instrumental resolution

Every spectrometer has a finite *resolution*. It can be described as a conditional probability $R(\omega|\omega')$ that a scattering event with an energy transfer of ω' is registered in the channel ω. Accordingly, a true, 'theoretical' spectrum $S^{\mathrm{th}}(\omega')$ gives rise to an observed, 'experimental' spectrum

$$S^{\mathrm{ex}}(\omega) = \int \mathrm{d}\omega'\, R(\omega|\omega') S^{\mathrm{th}}(\omega'). \tag{29}$$

In QENS, it is a good approximation that R depends only on the energy difference, $R(\omega|\omega') = R(\omega - \omega')$. Under this assumption, (29) is a convolution integral,

$$S^{\mathrm{ex}} = R \otimes S^{\mathrm{th}} \tag{30}$$

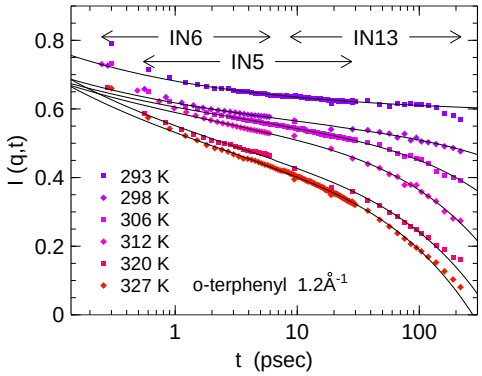


Fig. 3: *Intermediate scattering function of the glass-forming liquid ortho-terphenyl, measured on three different spectrometers of the ILL, and combined after Fourier deconvolution. Solid lines are fits with a mode-coupling scaling function. Redrawn from [5].*

for short, and R can be experimentally determined by measuring the spectrum of a material that is known to be a purely elastic scatterer, since $R \otimes \delta = R$.[4]

The resolution functions of time-of-flight and backscattering spectrometers are in a first approximation Gaussian. Typical scattering functions are qualitatively different in shape. When scaled for equal maximum and equal width at half maximum, a Lorentzian (and *a fortiori* a Kohlrausch-Williams-Watts function) has much broader wings than a Gaussian. In consequence, the onset of quasielastic scattering is regularly detected as additional scattering deep in the wings of the resolution function before any broadening is observed in the width at half maximum (Fig. 2). For this reason, in high-resolution neutron scattering the signal-to-noise ratio is a more important figure of merit than the nominal resolution width.

In principle, resolution effects can be removed from experimental data by Fourier deconvolution:

$$I^{\mathrm{th}}(t) = \frac{I^{\mathrm{ex}}(t)}{\tilde{R}(t)}. \tag{31}$$

The number of independent t points is limited by the Nyquist sampling theorem. For most of these t, (31) results in the division of two small, noisy numbers. Therefore one must introduce a cut-off time, restricting I^{th} to a relatively small number of short-time data points.

so that one has to restrict the computation to a relatively small number of short-time data points. This loss of information is normally not acceptable; instead of deconvoluting experimental data, it is preferable to fit the measured data S^{ex} with a theoretical function S^{th} that has been numerically convolved with the measured resolution R (or a smoothened model thereof). However, explicit Fourier deconvolution is attractive for combining spectral measurements from different spectrometers (Fig. 3) or for comparing neutron scattering with molecular dynamics simulations.

[4] Some phonon scattering can be tolerated in the resolution measurement, since it occurs mostly outside the energy window relevant for QENS. In practice, the resolution measurement is usually done either on vanadium (which is a perfectly incoherent scatterer so that the same measurement can also be used for detector calibration) or on the sample at low temperature (which has the advantage of being as close as possible to the conditions of the production measurements).

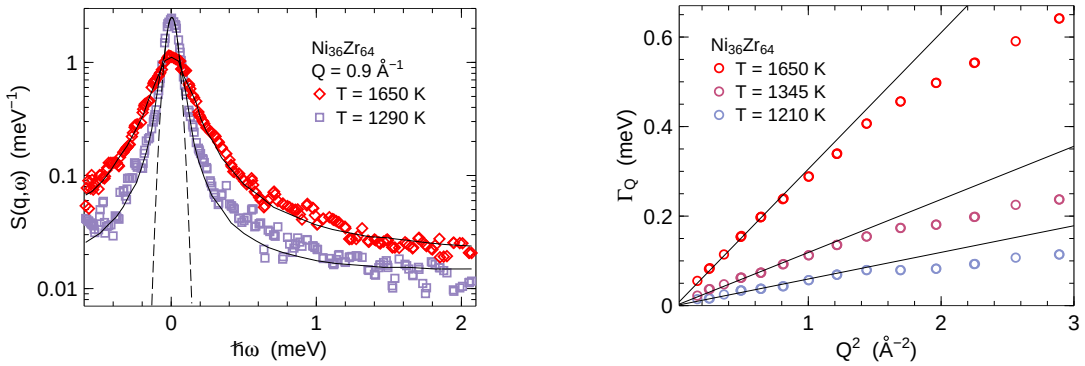


Fig. 4: *Atomic dynamics of liquid Zr-Ni, investigated with the time-of-flight spectrometer TOFTOF at FRM II, redrawn from [6]. (a) Selected spectra with Lorentzian fits (34). The dashed Gaussian with a FWHM of 95 μeV is an idealized representation of the instrumental resolution. (b) Linewidths Γ_Q, from Lorentzian fits to the measured spectra. For small Q, the scattering is dominated by incoherent contributions from Ni. Therefore, the initial slope of Γ_Q vs. Q^2 yields the self-diffusion coefficient of Ni.*

3 Application Examples

3.1 Diffusion

Diffusion has been discussed in much detail in lecture B3. To derive the incoherent scattering function for a diffusing particle, we equate the space-time probability distribution function $P(\mathbf{r}, t)$ introduced in section 2.2 of B3 with the classical approximation to the self-correlation function introduced in Eq. (9) of the present lecture. We can then immediately copy the solution of the standard diffusion equation,

$$G_{\mathrm{s}}^{\mathrm{cl}}(\mathbf{r}, t) = \frac{1}{(4\pi D t)^{3/2}} \mathrm{e}^{-r^2/4Dt}. \tag{32}$$

By Fourier transform, we find the intermediate scattering function

$$I(Q, t) = \exp(-DQ^2 t) = \tilde{\mathcal{L}}(t; DQ^2), \tag{33}$$

and by looking up (26) we obtain the scattering function

$$S(Q, \omega) = \mathcal{L}(\omega; DQ^2). \tag{34}$$

On a time-of-flight spectrometer, with experimental scales of the order $Q \sim \text{Å}^{-1}$ and $\hbar\omega \sim 0.1 \ldots 10$ meV, one can resolve diffusion coefficients D of the order $10^{-10} \ldots 10^{-8}$ m^2/s.

The straightforward determination of D from Lorentzian fits (34) works best in simple atomic systems. Recent examples are provided by metallic melts, which can be studied under very clean experimental conditions using electromagnetic levitation (Fig. 4). Results improve significantly upon macroscopic laboratory measurements that suffer from convective contributions.

In molecular liquids the applicability of (34) is not ascertained a priori because the atomic motion seen by neutron scattering is a superposition of molecular translation, molecular rotation, and innermolecular vibrations and rearrangements. This has been demonstrated very clearly in a systematic study of alkanes $\mathrm{C}_n\mathrm{H}_{2n+2}$ (Fig.5).

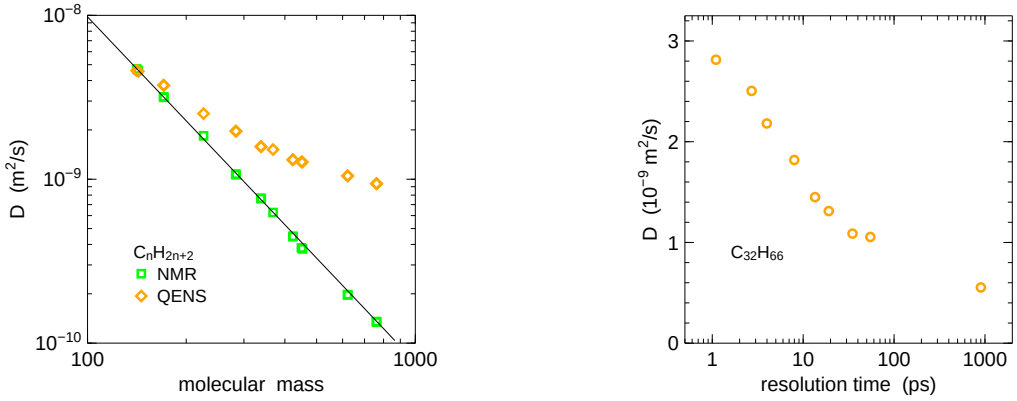


Fig. 5: *Hydrogen dynamics in n-alkanes. (a) Diffusion coefficient D, determined by pulsed-field gradient NMR and by neutron scattering using TOFTOF at FRM II. The Data points agree only for the shortest molecule ($n = 8$). In longer chains, QENS deviates from NMR because it measures not only center-of-mass molecular translation, but also rotation and innermolecular motion. Redrawn from [7]. (b) In an exemplary series of TOFTOF measurements with different instrumental resolutions (expressed through a resolution time), oversimplified fits have been shown to result in a resolution dependence of the apparent diffusion coefficients. Redrawn from [8].*

3.2 Jumps between two positions

In the simplest jump model, we consider jumps of a proton between two positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. This two-site jump model has only few applications, but it allows us to introduce concepts that come to fruition in the study of molecular rotation.

The probability $p(\boldsymbol{r}, t)$ of finding the proton at time t at site $\boldsymbol{r}$ obeys the rate equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} p(\boldsymbol{r}_1,t) \\ p(\boldsymbol{r}_2,t) \end{pmatrix} = -\begin{pmatrix} \lambda_1 & -\lambda_2 \\ -\lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} p(\boldsymbol{r}_1,t) \\ p(\boldsymbol{r}_2,t) \end{pmatrix} \tag{35}$$

with transition rates λ_n. The matrix has the eigenvalues 0 and $\Gamma := \lambda_1 + \lambda_2$, so that the rate equation is solved by $p(\boldsymbol{r}_n, t) = a_n + b_n \exp(-\Gamma t)$. In the long-time limit $t \to \infty$, the occupation ratio must be $p_1/p_2 = \lambda_2/\lambda_1$ to satisfy $\mathrm{d}p/\mathrm{d}t = 0$. Combined with the normalization condition $\sum_n p(\boldsymbol{r}_n, t) = 1$, we find

$$a_1 = p(\boldsymbol{r}_1, \infty) = \frac{\lambda_2}{\Gamma}, \quad a_2 = p(\boldsymbol{r}_2, \infty) = \frac{\lambda_1}{\Gamma}. \tag{36}$$

Imposing the initial condition $p(\boldsymbol{r}_1, 0) = 1$, we compute the conditional probabilities

$$\begin{aligned} p(\boldsymbol{r}_1, t|\boldsymbol{r}_1, 0) &= a_1 + a_2 \exp(-\Gamma t), \\ p(\boldsymbol{r}_2, t|\boldsymbol{r}_1, 0) &= a_2 \left(1 - \exp(-\Gamma t)\right), \end{aligned} \tag{37}$$

and similarly for $p(\boldsymbol{r}_2, 0) = 1$. Using the equilibrium occupation probabilities (36), and introducing the jump vector $\boldsymbol{d} := \boldsymbol{r}_2 - \boldsymbol{r}_1$, we obtain the intermediate self correlation function

$$\begin{aligned} I(\boldsymbol{Q}, t) &= \langle \mathrm{e}^{i\boldsymbol{Q}\boldsymbol{r}(t)} \mathrm{e}^{-i\boldsymbol{Q}\boldsymbol{r}(0)} \rangle \\ &= p(\boldsymbol{r}_1, \infty)\left[p(\boldsymbol{r}_1, t|\boldsymbol{r}_1, 0) + p(\boldsymbol{r}_2, t|\boldsymbol{r}_1, 0)\mathrm{e}^{i\boldsymbol{Q}\boldsymbol{d}}\right] + \\ &\quad p(\boldsymbol{r}_2, \infty)\left[p(\boldsymbol{r}_2, t|\boldsymbol{r}_1, 0) + p(\boldsymbol{r}_1, t|\boldsymbol{r}_1, 0)\mathrm{e}^{-i\boldsymbol{Q}\boldsymbol{d}}\right]. \end{aligned} \tag{38}$$

Regrouping terms, abbreviating

$$\begin{aligned} A_0(\boldsymbol{Q}) &:= a_1^2 + a_2^2 + 2a_1a_2 \cos \boldsymbol{Qd}, \\ A_1(\boldsymbol{Q}) &:= 2a_1a_2(1 - \cos \boldsymbol{Qd}), \end{aligned} \tag{39}$$

and evaluating the Fourier transform of $\exp(-\Gamma t)$, we get the incoherent scattering function

$$S(\boldsymbol{Q}, \omega) = A_0(\boldsymbol{Q})\delta(\omega) + A_1(\boldsymbol{Q})\mathcal{L}(\omega; \Gamma), \tag{40}$$

which consists of an elastic line and a quasielastic component, the latter having standard Lorentzian shape (26). As discussed above, the presence of an elastic line is characteristic for *localized motion*: Since the jumping proton is confined to a finite region in space, its self correlation function never decays to zero.

For measurements performed on powder samples, we average over the orientations of $\boldsymbol{d}$,

$$\overline{\cos \boldsymbol{Qd}} = \frac{1}{4\pi} \int_0^\pi \mathrm{d}\vartheta \, 2\pi \sin\vartheta \, \cos(Qd \cos\vartheta) = \frac{\sin Qd}{Qd} = j_0(Qd), \tag{41}$$

where the last equation introduces a customary abbreviation, the spherical Bessel function j_0. In the simplest case, for jumps between two equivalent positions, we have $\lambda_1 = \lambda_2$ and $a_1 = a_2 = 1/2$, so that the powder-averaged amplitudes take the form

$$\begin{aligned} A_0(Q) &= (1 + j_0(Qd))/2, \\ A_1(Q) &= (1 - j_0(Qd))/2. \end{aligned} \tag{42}$$

3.3 Rotational jump diffusion

To discuss the rotation of molecules or molecular side groups in solids we choose a simple and practically important example, a methyl group R–CH_3. We consider the group as stiff (CH bond length $d = 1.097 \pm 0.004$ Å, HCH angle $\theta = 106.5 \pm 1.5°$). The only degree of freedom is the rotation around the bond that connects the methyl group to the remainder R of the molecule. This R–C bond coincides with the symmetry axis of the CH_3 group. The corresponding moment of inertia is

$$I = \sum m d_\perp^2 = 2md^2(1 - \cos\theta). \tag{43}$$

The rotational motion can be described by a wave function ψ that depends on one single coordinate, the rotation angle ϕ. The Schrödinger equation is

$$\left\{ B\frac{\partial^2}{\partial\phi^2} - V(\phi) + E \right\} \psi(\phi) = 0 \tag{44}$$

with the *rotational constant*

$$B := \frac{\hbar^2}{2I} = 670 \ \mu\mathrm{eV}. \tag{45}$$

For free rotation ($V = 0$), solutions that possess the requested periodicity are sine and cosine functions of argument $J\phi$, with integer J. Accordingly, the energy levels are $E = BJ^2$.

In condensed matter, however, the potential V caused by the local environment cannot be neglected. Due to the symmetry of the CH_3 group, the Fourier expansion of $V(\phi)$ contains

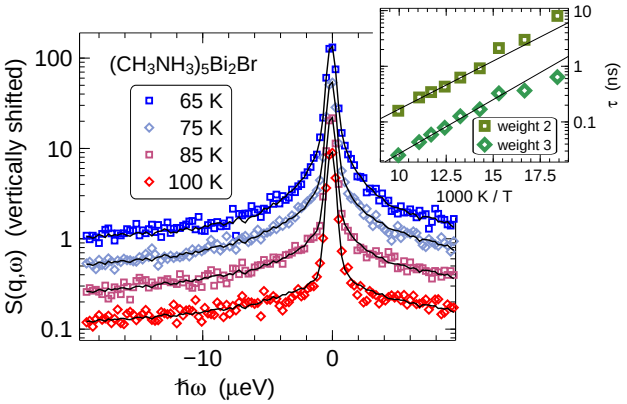


Fig. 6: *Backscattering spectra of $(CH_3NH_3)_5Bi_2Br$, measured on SPHERES [9]. The five methylammonium cations fall into two different categories: at room temperature, two of them are ordered, three are disordered [10]. Therefore, we fitted the spectra with two Lorentzians with an amplitude ratio of 2:3. The resulting relaxation times have an Arrhenius temperature dependence, shown in the inset.*

only sine and cosine functions with argument $3m\phi$, with integer m. In most applications, it is sufficient to retain only one term,

$$V(\phi) \doteq V_3 \cos(3\phi). \tag{46}$$

The strength of the potential can then be expressed by the dimensionless number V_3/B. In the following we specialize to the case of a *strong potential*, $V_3/B \gg 10$, which is by far the most frequent one.

In a strong potential of form (46), the CH_3 group has three preferential orientations, separated by potential walls. The motion of the CH_3 group consists mainly of small excursions from the preferred orientations, called *librations*. Quantum-mechanically, they are zero-point oscillations in an approximately harmonic potential.

Orientational motion can be approximated as thermally activated *jump diffusion* between equivalent equilibrium positions. For instance, to compute incoherent scattering from a rotating methyl (CH_3) group, it can be sufficient to consider 120° jumps between three equivalent rest positions on a circle of radius r.

This requires only a little extension of the two-site jump model introduced above. The transition matrix in the rate equation takes the form

$$\begin{pmatrix} 2\lambda & -\lambda & -\lambda \\ -\lambda & 2\lambda & -\lambda \\ -\lambda & -\lambda & 2\lambda \end{pmatrix}, \tag{47}$$

which has the eigenvalues 0, 3λ, 3λ. Thanks to the degeneracy of the nonzero eigenvalue, the scattering law retains the simple form (40), with $\Gamma = 3\lambda$, and with amplitudes

$$\begin{aligned} A_0(Q) &= (1 + 2j_0(Qr\sqrt{3}))/3, \\ A_1(Q) &= (2 - 2j_0(Qd\sqrt{3}))/3. \end{aligned} \tag{48}$$

This model has proven successful in a huge number of experiments; Fig. 6 shows an arbitrarily chosen recent example.

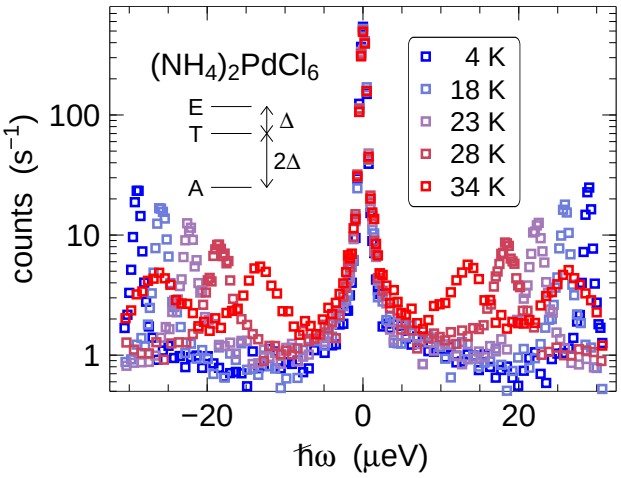


Fig. 7: *Backscattering spectra of the perovskite $(NH_4)_2PdCl_6$, measured on SPHERES [12]. The inelastic lines are due to rotational tunneling of the NH_4^+ group. The energy-level scheme holds for tetrahedral symmetry at the NH_4^+ sites if the probability of 180° jumps is negligible compared to 120° jumps [13]. Only the E↔T transition is observed, except at 34 K where also the A↔T transition fits into the experimental energy range.*

However, Eq. (40) must be modified if the rotational potential has a C_2 symmetry. Then there are six equivalent equilibrium positions, connected by 60° jumps, so that the transition matrix is of rank 6. After some computation it is found to have three different non-zero eigenvalues Γ_μ. In such a situation, the inelastic part of scattering law no longer factorises into a Q dependent and a ω dependent function. Instead, one has a sum of Lorentzians of different widths:

$$S(\boldsymbol{Q},\omega) = A_0(\boldsymbol{Q})\delta(\omega) + \sum_\mu A_\mu(\boldsymbol{Q})\mathcal{L}(\omega;\Gamma_\mu). \tag{49}$$

This equation holds quite generally for systems described by a rate equation of the form (35) with an arbitrary, symmetric transition matrix. In particular, it holds for rotational jump diffusion of molecules that have more than one axis of rotation [11].

3.4 Rotational tunneling

At low temperatures, almost exclusively the vibrational ground state is occupied. Yet reorientational motion beyond librations is possible by means of quantum mechanical tunneling: The wave functions of the three localised *pocket states* ψ_m ($m = 1,2,3$) have nonzero overlap. Therefore, the eigenstates are a linear combination of pocket states.[5] Periodicity and threefold symmetry allow three such combinations: a plain additive one

$$\psi_1 + \psi_2 + \psi_3, \tag{50}$$

and two superpositions with phase rotations

$$\psi_1 + \mathrm{e}^{\pm i2\pi/3}\psi_2 + \mathrm{e}^{\pm i4\pi/3}\psi_3. \tag{51}$$

[5]This is an extremely simplified outline of the theory. In a serious treatment, to get all symmetry requirements right, one must also take into account the nuclear spins of the H atoms [14].

In the language of group theory, state (50) has symmetry A, the degenerate states (51) are labelled E^a, E^b. It is found that A is the ground state. The *tunneling splitting* $\hbar\Omega_\mathrm{t}$ between the states A and E is determined by the overlap integral $\langle\psi_m|V|\psi_n\rangle$ ($m \neq n$), which depends exponentially on the height of the potential wall. Experiments that detect tunneling transitions provide therefore a very sensitive probe of the rotational potential; conversely, if the potential is not accurately known, it is almost impossible to predict whether a tunneling transition will show up in a given experimental energy range.

In neutron scattering, a tunneling transition appears as a pair of inelastic peaks at $\pm\hbar\Omega_\mathrm{t}$. The spectral shape of these peaks is well described by Lorentzians $\mathcal{L}(\omega \pm \Omega_\mathrm{t}; \Gamma)$. With rising temperatures, the occupancy of excited vibrational levels increase. This facilitates transitions between A and E sublevels and results in a decrease of $\hbar\Omega_\mathrm{t}$ and an increase of the line width Γ.

Upon further temperature increase, thermal motion of neighbouring molecules causes so strong potential fluctuations that the picture of quantum tunneling is no longer applicable. Instead, the motion between different pocket states must be described as *stochastic jump diffusion*, as exposed above.

For systems with more than one rotational axis, group theory is used to identify eigenstate symmetries. For instance for tetrahedral NH_4^+ ions in tetrahedral cages, three energy levels are found, with a transition energy ratio of $\Omega_\mathrm{TA} = 2\Omega_\mathrm{ET}$ (Fig. 7).

References

[1] N. W. Mermin, J. Math. Phys. **7**, 1038 (1966).

[2] W. Doster, S. Cusack and W. Petry, Nature **337**, 754 (1989).

[3] W. Doster, Eur. Biophys. J. **37**, 591 (2008).

[4] W. Doster, S. Busch, A. M. Gaspar, M. S. Appavou, J. Wuttke and H. Scheer, Phys. Rev. Lett. **104**, 098101 (2010).

[5] J. Wuttke, M. Kiebel, E. Bartsch, F. Fujara, W. Petry and H. Sillescu, Z. Phys. B **91**, 357 (1993).

[6] D. Holland-Moritz, S. Stüber, H. Hartmann, T. Unruh, T. Hansen and A. Meyer, Phys. Rev. B **79**, 064204 (2009).

[7] C. Smuda, S. Busch, G. Gemmecker and T. Unruh, J. Chem. Phys. **129**, 014513 (2008).

[8] T. Unruh, C. Smuda, S. Busch, J. Neuhaus and W. Petry, J. Chem. Phys. **129**, 121106 (2008).

[9] A. Piecha, J. Wuttke, R. Jakubas and G. Bator, JCNS experimental report 3961 (2010).

[10] J. Matuszewski, R. Jakubas, L. Sobczyk and T. Głowiak, Acta Cryst. C **46**, 1385 (1990).

[11] M. Bée, *Quasielastic Neutron Scattering*, Adam Hilger: Bristol (1988).

[12] J. Wuttke, JCNS experimental report 2888 (2010).

[13] A. Hüller, Phys. Rev. B **16**, 1844 (1977).

[14] W. Press, *Single-particle rotation in molecular crystals* (Springer tracts in modern physics 92), Springer: Berlin (1981).

D 8 High Resolution Spectroscopy

G. Meier[1] and M. Monkenbusch[2]
[1]Institute of Complex Systems, ICS-3
[2]Jülich Centre for Neutron Science 1
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

In this lecture we are concerned with the extensions of the measurement capabilities in neutron and x-ray scattering beyond that what was known in the seventies regarding neutron scattering and in the nineties regarding x-ray scattering. In table 1 a sketch of the time/energy versus scattering vector/correlation length plane is shown. In this diagram the two relevant techniques under study in this presentation are neutron spin echo spectroscopy (NSE) and X-ray photon correlation spectroscopy (XPCS). The former method extends the usually accessible time/energy range of neutron scattering towards much longer times or lower energies, whereas the x-ray correlation spectroscopy extends the usually accessible Q-range of dynamic light scattering toward much higher Q-values. Since both methods are concerned with an extension of what was known before they are therefore named high resolution techniques, although they deserve this allocation for physically different reasons.

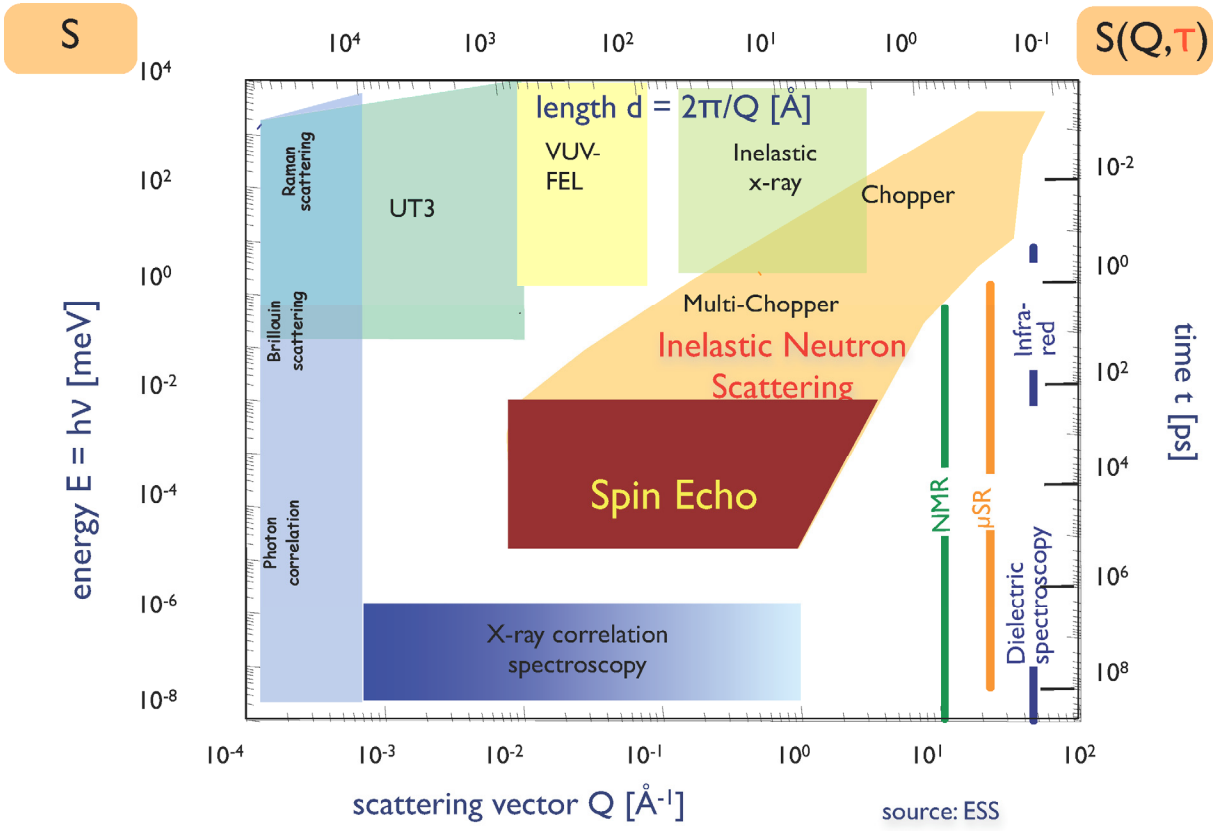


Fig. 1: *Sketch of various spectroscopic techniques with regard to the accessible ranges in the scattering vector Q and correlation time τ.*

X-ray photon correlation spectroscopy, XPCS, similar to dynamic light scattering, DLS, analyses the intensity correlations of scattered electromagnetic radiation. However, the contrast results from physically different properties: index of refraction differences in DLS in contrast to differences in electron density for XPCS. The other mayor difference is the accessible Q-range, mainly due to the used wavelength. The access to very short correlation times is restricted in XPCS, because either the storage ring from the synchrotron has a time structure or the scattering cross section is so small that practically no decent counting statistics can be achieved.

The spin-echo principle stems from the Lamor-precession of polarized neutrons within a magnetic field. In general it is realized via a time of flight spectrometer set-up, in which a time of flight difference for a fixed distance is measured before and after the neutron has interacted with a sample. The probe to do so is the polarisation of the neutrons. According to Fig.1, correlation time values up to several hundreds of nanoseconds are available. Note that the spatial resolution of the scattering experiment is in the nanometer range, which means that a time range of e.g. 100 ns corresponds to effective molecular motion velocities of 1 nm/100 ns = 1 cm/s. This may be compared to the typical neutron velocity of 200-1000 m/s used in these types of experiments. Hence it becomes obvious why NSE is called a high resolution spectroscopic method. The basic idea for this neutron spectroscopy was brought forward by F. Mezei in 1972 [1].

All mentioned scattering techniques aim to measure a scattering function that is closely related to the density-density correlations

$$S_{i,j}\left(\vec{Q},\omega\right)=\frac{1}{2\pi\hbar}\int dt\int d\vec{r}\exp\left(-i\omega t-i\vec{Q}\cdot\vec{r}\left\langle\rho_i\left(\vec{r}'-\vec{r},t\right)\rho_j\left(\vec{r}',t'\right)\right\rangle\right) \qquad (1)$$

where $\rho_i(\vec{r},t)$ denotes the density of species i at position $\vec{r}$ and time t. The intermediate scattering function is the Fourier transform of S(Q,ω) in the frequency-time domain:

$$S_{i,j}\left(\vec{Q},t\right)=\hbar\int d\omega\exp\left(i\omega t\right)S_{i,j}\left(\vec{Q},\omega\right) \qquad (2)$$

The observed time dependence of S(Q,ω), respectively of the density correlations results from thermal fluctuations (Brownian motions) that are driven by random uncorrelated forces due to coupling to the heat bath.

2 Principles of Neutron Spin-Echo Spectroscopy

The scattering intensity observed in a neutron scattering experiment depends besides trivial factors like the sample amount, transmission and incoming beam intensity on the double differential cross section:

$$\frac{d^2\sigma}{d\Omega dE'} = C\,\frac{k'}{k}\sum_{i,j} b_i b_j\, S_{i,j}(\vec{Q},\omega) \tag{3}$$

where k', k is the modulus of the scattered and incoming neutrons respectively. C is proportional to the amount of sample and b_i is the scattering length of atoms of type i. The energy transfer during scattering is $\hbar\omega = (E - E')$, with E, E' the energies of the incoming and scattered neutrons respectively. The variables of the scattering function depend on $\vec{k}$, $\vec{k}'$; $\vec{Q} = \vec{k} - \vec{k}'$ is the momentum transfer and

$$\hbar\omega = (\hbar^2/2m_n)(k^2 - k'^2) \tag{4}$$

the energy transfer that occurred during the scattering process. $\vec{k}$ relates linearly to $\vec{v}$ (momentum $\vec{p} = m_n\vec{v}$) as to 1/wavelength:

$$\vec{v}m_n = \hbar\vec{k}\,\frac{2\pi}{\lambda} = k \tag{5}$$

It is possible to determine the energy transfer (Eq.4) by measurement of the velocity difference $v' - v$ of the scattered neutrons compared to the incoming neutrons. For large relative velocity changes $\Delta v/v > 1\%$ this measurement could be done by comparing the average velocity of the incoming beam with the velocity distribution of the scattered neutrons. However, if very small differences have to be detected -as it is necessary for the analysis of "soft matter" dynamics- the preparation of a sufficiently narrow velocity distribution of the incoming beam from the continuous (Maxwellian) velocity spectrum emitted by the moderator of the neutron source would leave only a very small and insufficient number of neutrons in the incoming beam.

This situation would immediately improve, if it would be possible to equip each neutron with an individual stop watch which could be read in a way that the run time difference between test tracks before and after the sample is obtained at detection. If this "stop watch" had a sufficient time resolution it would be possible to observe very small velocity changes even if a beam with a wide range of initial neutron velocities is used. This allows to escape the intensity trap.

With some restrictions it is indeed possible to use the neutron spin directions as kind of individual stop watch pointers. The clockwork of this watch is then effected by the precession of the neutron spins in an external magnetic field. The restrictions affecting application are caused by the fact that the "spin-stop watch" can only be read up to an

unknown integer number of complete precession turns. Using a polarised incoming beam and a sequence of subsequent spin rotations, the reading is performed by the cosine type transmission function of an analyser and yields only ensemble averages and not individual rotation angles. The intensity at the detector is modulated accordingly. The encountered Δv values are distributed according to $S(Q,\omega \approx k\Delta v)$. The detector signal is then proportional to the integral of precession-angle-cosine modulated intensity contributions with weight according to the distribution of Δv implied by $S(Q,\omega)$. I.e. it is proportional to the cosine-Fourier transform of $S(Q,\omega)$, the intermediate scattering function $S(Q,t)$.

Before we start a more detailed derivation a concrete setup of an NSE spectrometer (see Fig.2) is presented [1,2,3]. A longitudinally polarised neutron beam enters the spectrometer and hits the first so-called $\pi/2$-flipper where the spin is rotated such that on exit it is orthogonal to the longitudinal magnetic field along the following precession path.

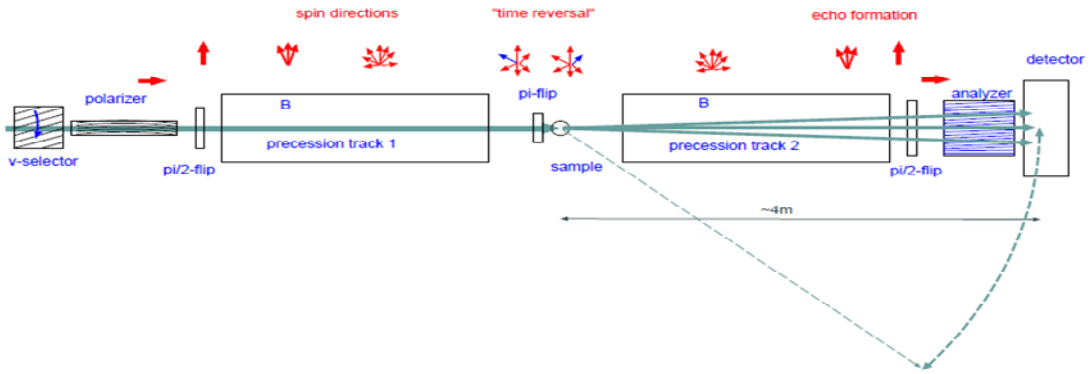


Fig. 2: *The generic setup of a neutron spin-echo spectrometer. Neutrons from a reactor source enter from left, are coarsely monochromatised by a mechanical velocity selector and polarised by reflection from magnetic multilayers. Then with the first π/2- flipper they enter the proper NSE spectrometer and the neutron precession starts. Velocity coding into precession angle is affected by a magnetic field in precession track 1, the π-flipper close to the sample rotates the spin ensemble by 180° and thereby causes an effective "time reversal". Elastic scattering at the sample without velocity change of the neutron then allows the reassembly of all precession angles into a spin-echo. Inelastic scattering at the sample renders this effect less efficient and leads to a reduction of neutron polarisation in the echo condition. This contains the desired information.*

That defines the start of the "spin stop watch", immediately after the flipper a precession of the spins around the axial magnetic field begins. The precession frequency increases during the approach of the center of the main precession solenoid where it reaches its maximum of up to a few MHz. The accumulation of precession angle continues -with decreasing frequency- until the neutrons reach the π-flipper close to the sample (S). The total precession angle at that point is:

$$\Psi = \frac{\gamma}{v}\int_l |B|\,dl = n2\pi + \alpha \tag{6}$$

where $\gamma = 2\pi \times 2913.06598 \times 10^4 \mathrm{s}^{-1}/Tesla$ is the gyromagnetic ratio of the neutrons and $|B|$ is the modulus of the magnetic induction along the path l.

Close to the sample (ideal: at the sample position) the so-called π-flipper is located, it rotates the spins by 180^0 around a vertical axis, thereby the total precession angle is transformed to $\Psi_1 = n2\pi + \alpha \rightarrow n2\pi - \alpha$. The parameters n and α are -according to Eq.6 - extremely dependent on velocity and therefore very different for different neutrons in a beam with finite width of the wavelength distribution. As a consequence the spin vectors at the sample position (π-flipper) are evenly distributed on a disc orthogonal to the field direction. If no velocity change occurs during scattering at the sample (elastic scattering) each neutron enters the secondary arm of the spectrometer with unchanged velocity. The precession field and path length of the second arm exactly match the corresponding elements of the primary arm before the sample (π-flipper). Accordingly the precession accumulated in the secondary arm is $\Psi_2 = n2\pi + \alpha$ and the total precession angle at the second $\pi/2$-flippers is $\Psi_1 + \Psi_2 = (n+n)2\pi - \alpha + \alpha = 2n2\pi$. I.e. all spins -irrespective of their initial velocity- reassemble at the same vertical position they had at the start point, the rotation imposed by the second π-flipper converts this back to the initial longitudinal polarisation that is fully restored. The flippers limit the two race tracks and realise "start", "time reversal" and "stop" of the "spin stop watches". The second $\pi/2$-flipper is the last element used to manipulate the spins it converts the average precession angle to a longitudinal polarisation component. Since the field after the second $\pi/2$-flipper is again longitudinal, further precessions do not influence the analysed longitudinal polarisation component (the stop watch is stopped!). The analyser consists -like the polariser- of magnetic multilayer mirrors which only reflect neutrons of one longitudinal spin state into the detector. After ensemble averaging this means that the count rate at the detector is proportional to $(1 \pm \cos(\overline{\Psi}))/2$, where $\overline{\Psi}$ is the expectation value of the angle between spin and axial direction.

In the following the relation between detector signal and $S(Q,\omega)$ is derived in terms of mathematical expressions. First the field integrals along the primary and secondary paths of precession are defined:

$$J_1 = \int_{l((\pi/2)_1)}^{l(\pi)} |B|\,dl \tag{7}$$

$$J_2 = \int_{l(\pi)}^{l((\pi/2)_2)} |B|\,dl \tag{8}$$

for a symmetric setup $J_1 = J_2$; $l(\pi,(\pi/2)_{1,2})$ denotes the positions of the corresponding flippers. The precession angle accumulated on a path is

$$\Psi_i = \frac{\gamma J_i}{v} \tag{9}$$

where v is the neutron velocity (typically several 100 m/s). Because the π-flipper inverts the sign of Ψ_1 (i.e. of the part $\Psi_1 \mathrm{m}od\ 2\pi$), a total precession angle of

$$\Psi_{1,2} = -\frac{\gamma J_1}{v} + \frac{\gamma J_2}{v+\Delta v} + 4n\pi \tag{10}$$

results, where Δv is the velocity change of the neutron during scattering and n is an integer. The transmission function of an (ideal) analyser is

$$T_a = \frac{1}{2}[1+\cos(-\frac{\gamma J_1}{v} + \frac{\gamma J_2}{v+\Delta v})] \tag{11}$$

From that the detector intensity

$$I = \eta S(Q) \iint \frac{1}{2}[1 \pm \cos(-\frac{\gamma J_1}{v} + \frac{\gamma J_2}{v+\Delta v})] w_\omega(\Delta v) w_\lambda(v)\, d\Delta v\, dv \tag{12}$$

results, where η is an irrelevant calibration factor and w_ω respectively. w_λ are normalized distribution functions. w_ω represents the spectrum of the sample as found in the scattering function and w_λ takes account for the fact that the NSE spectrometers usually are operated with a broad incoming wavelength distribution ($\Delta\lambda_{FWHM}/\lambda = 10\cdots 20\%$). Observing the linear dependencies of k , $1/\lambda$ and v and series expansion of the squares in Eq.4 and insertion into Eq.12 leads to:

$$I = \eta \iint \frac{1}{2}[1 \pm \cos(-\frac{\gamma J_1}{(h/m_n)\lambda^{-1}} + \frac{\gamma J_2}{(h/m_n)\lambda^{-1} + \lambda\omega/2\pi})] S(Q,\omega) w_\lambda(\lambda)\, d\omega\, d\lambda \tag{13}$$

After some more mathematics observing the distribution of incoming wavelengths and the distribution of deviations of the field integrals from their nominal value (inhomogeneity) the following approximate expression for the detected intensity is obtained:

$$I_{Det} \propto \frac{1}{2}\left[S(Q) \pm \frac{\exp(-\Psi^2)}{N} \cos(\delta\gamma \frac{m_n}{h}\lambda) \int \cos(J\lambda^3\gamma \frac{m_n^2}{2\pi h^2}\omega)\, S(Q,\omega) d\omega \right] \tag{14}$$

$\delta = J_1 - J_2$ is the asymmetry of the field integral between both spectrometer arms. Under the assumption that the incoming wavelength distribution and the distribution of field integrals

around their nominal values are Gaussians, the following explicit relations hold: $N^2 = 1 + 4\Sigma^2\Lambda^2\gamma^2 m_n^2/h^2 \simeq 1$ and $\Psi^2 = [\Sigma^2\lambda_0^2 + \Lambda^2\delta^2]\gamma^2 m_n^2/h^2/N^2$. $\Sigma^2 = \langle \Delta J^2 \rangle/2$ relates to the field integral inhomogeniety and $\Lambda = \Delta_{FWHM}\lambda/4/\sqrt{\ln 2}$ to the full width at half maximum (FWHM) of the incoming wavelength distribution. For a given 10% wavelength width and $\lambda_0 = 1nm$ and $\Sigma = 2\times10^{-6}\mathrm{T}m$ the resulting resolution factor is $R = \exp(-\Psi^2) = 0.43$.

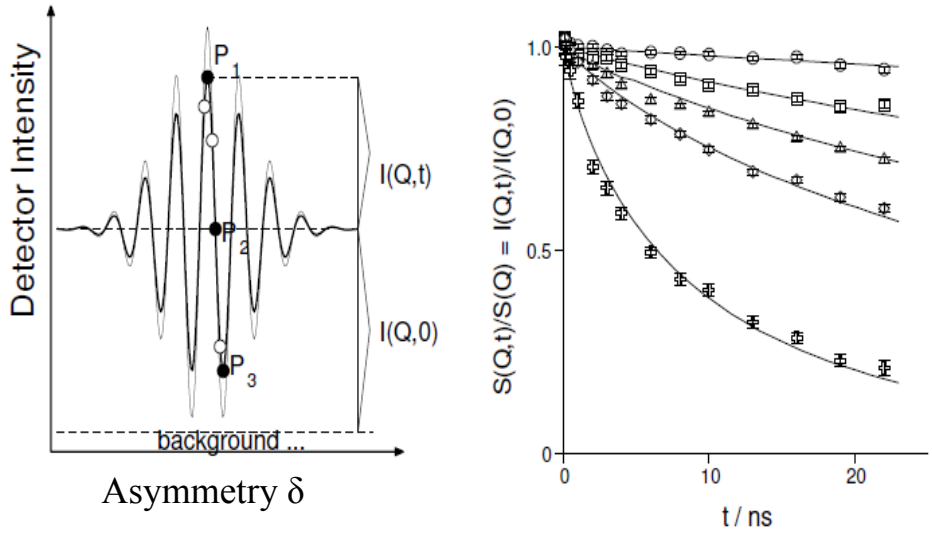


Fig. 3: *The left side indicates the detector intensity as obtained if the symmetry of the two precession fields is varied, the width of the echo oscillation group is inversely proportional to the width of the incoming wavelength (here as example more than 20% FWHM). The amplitude which may be determined by measuring 3 or more points and eventually by fitting the echo signal to the set of points ($\{P_1, P_2, \cdots, P_n\}$) compared to the maximum possible amplitude contains the spectroscopic signal which is obtained by NSE. The right side shows a typical set of results as obtained from a polymer melt with single chain contrast (i.e. a few protonated chains in a deuterated melt). Each point in that plot corresponds to one elementary amplitude determination. The various Fourier times result from different magnetic precession field settings and the various curves correspond to Q-values 0.05, 0.08, 0.12, 0.19 $\mathrm{Å}^{-1}$. The lines are the predictions of the Rouse model[4,5,6].*

In an experiment , see Fig.3, the asymmetry δ is varied to determine the amplitude, A, of the first cosine factor in Eq.14:

$$A = \pm \underbrace{\exp(-\Psi^2)/N}_{=R} \int \cos(\underbrace{J\lambda^3\gamma \frac{m_n^2}{2\pi h^2}}_{=t}\ \omega)\, S(Q,\omega) d\omega \qquad (15)$$

i.e. up to the resolution R the amplitude A is proportional to $S(Q,t)$. The resolution factor R is determined by measuring a reference sample with purely elastic scattering.

2.1 Example for Spin Echo Spectroscopy

The example presented here deals with the determination of the bending modulus of a microemulsion [7]. Microemulsions are thermodynamically stable mixtures of water and oil mediated by a surfactant. The typical structural sizes d of oil/water domains are in the order of d=20nm. Various structures can be obtained ranging from bicontinuous to lamellar microemulsions, depending on the interface properties and eventually their interaction. Different approaches have been developed for describing the physical properties of such microemulsions. Ginzburg-Landau models lead to a fair description of the scattering properties of a bicontinuous microemulsion [8,9].

$$S(Q) \propto \left(Q^4 - 2\left(Q_0^2 - \xi^{-2}\right)Q^2 + \left(Q_0^2 - \xi^{-2}\right)^2 \right)^{-1} \tag{16}$$

with $Q_0 = 2\pi/d$ and ξ the correlation length. Eq.16 matches the scattering from the scattering length contrast between oil and water in the wave vector region between $Q < \delta Q_0$ ($\delta \approx 2$-4). On the other hand a description of the free energy in terms of interface bending contributions is commonly used, where the Helfrich free energy, $F_{elastic}$, is expressed as

$$F_{el} = \int dS \left[\frac{\kappa}{2}\left(c_1 + c_2 - 2c_0\right)^2 + \bar{\kappa} c_1 c_2 \right] \tag{17}$$

with $c_i = 1/R_i$ the local values of the principal curvature, c_0 the spontaneous curvature (here about zero) and $\kappa, \bar{\kappa}$ the bending modulus and the saddle splay modulus respectively. According to the Gauss-Bonnet theorem the latter term only contributes to the free energy changes when the surface topology changes.

By matching the Ginzburg-Landau free energy expression that underlies Eq.16 to a Gaussian random field approach, a relation between the parameters d and ξ of Eq.16 and the renormalized bending modulus κ entering Eq.17 can be established [10] :

$$Q_0 \xi = \frac{64}{5\sqrt{3}} \frac{\kappa_R}{k_B T} \Theta \tag{18}$$

where Θ approaches 1 for large enough κ. The elastic bending energy F_{el} of the surfactant membrane controls the phase behaviour. To get a deeper insight experimentally into the structure of microemulsions small angle neutron scattering (SANS) is used. The characteristic distance d and the correlation length ξ can be obtained from fitting Eq.16 to SANS data. Basically d determines the position of the maximum in the static structure factor S(Q), and ξ its width. Complementarily, the membrane fluctuation dynamics as investigated with neutron spin echo spectroscopy (NSE) also carries information on the interface bending modulus.

NSE allows the analysis of dynamic correlations in the neutron scattered intensity in the relevant SANS Q-regime.

Neutron spin echo spectroscopy measures the intermediate scattering function S(Q,t) usually given in the normalized form as relaxation function S(Q,t)/S(Q). In our example presented here, NSE data were taken in the so-called film contrast, i.e. deuterated oil and heavy water. Only the surfactant was protonated to assure proper contrast of the membrane only. The dynamics of fluctuating membranes in a microemulsion can be described by the model of a fluctuating membrane in a viscous medium, as has been proposed by Zilman und Grenek [11]. For small deformations of a nearly planar membrane, the free energy of deformation of an undulating patch of membrane can be described in terms of the Helfrich-Hamiltonian [12]:

$$H = \frac{1}{2}\kappa \int d^2 r \left[\nabla_r^2 h\left(\vec{r}\right)\right]^2 = \frac{1}{2}\kappa \sum_{\vec{k}} k^4 h_{\vec{k}} h_{-\vec{k}} \tag{19}$$

where $h(\vec{r})$ denotes the amplitude of the undulation (the deviation of the membrane from a flat surface), and $h_{\vec{k}}$ its Fourier-components. The wave vector of the undulation modes is $\vec{k}$ and must not be confused with the scattering vector $\vec{Q}$.The Helfrich-Hamiltonian leads to the relaxation rate of $\left\langle h_{\vec{k}}(t) h_{-\vec{k}} \right\rangle_{Helfrich} = \left[k_B T / \left(\kappa k^4\right)\right] \exp\left[-\omega(k) t\right]$ with $\omega(k) = \kappa k^3 / 4\eta$. Besides the bending modulus also the average viscosity η of the surrounding medium determines the characteristic times.

Assuming an ensemble of independent , randomly oriented membrane patches that obey the above relation, Zilman and Granek deduced an approximation for S(Q,t) in terms of a stretched exponential function, with a stretching exponent β = 2/3 [11].

$$S(Q,t) \simeq S(Q) \exp\left[-\left(\Gamma_Q t\right)^{\beta}\right] \quad with \quad \Gamma_Q = 0.025 \gamma_k \left(\frac{k_B T}{\kappa}\right)^{0.5} \frac{k_B T}{\eta} Q^3 \tag{20}$$

where $\gamma\kappa \approx 1-(3k_BT/4\pi\kappa) \ln(Q\xi) \rightarrow 1$ for $\kappa \gg k_BT$. The functional form of Eq.20 describes the NSE-relaxation data, see Fig.4, and the parameter β approaches 2/3 for large $Q > 4Q_0$ and $\Gamma_Q \sim Q^3$, however, the absolute value for κ inferred from Eq.20 is seemingly wrong. Further refinements of the Zilman-Granek model have to be considered in order to get a consistent picture of the NSE data with respect to SANS results. First of all, the variation of the structural length scale by changing the surfactant volume fraction changes the reduced relaxation rate $\Gamma^* = \left(\Gamma_Q / Q^3\right)\eta(T)/k_BT$ significantly and next, the length scale of the structures considered that relates to the patch size of the model has to enter the final result. The full analysis of these refinements and to avoid further approximations that enter the Zilman-Granek model is shown in Ref. [7].

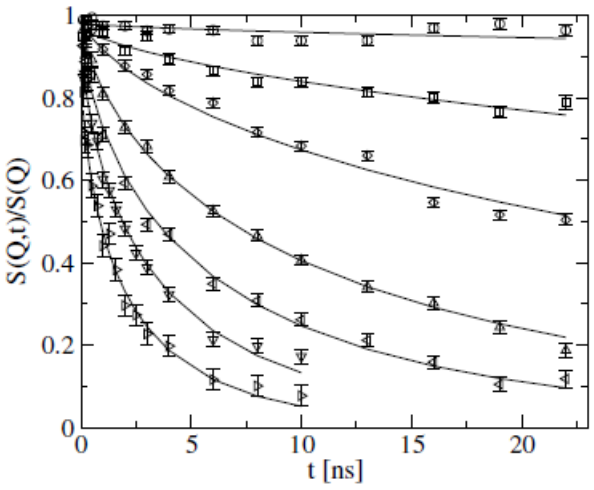


Fig. 4: *NSE spectra (Q = 0.05 – 0.24Å^{-1}) from a bicontinuous microemulsion fitted with a stretched exponential function where Γ_Q and β were separate fitting parameters for each curve. The β values approach 2/3 in the limit of high Q and $\Gamma_Q \sim Q^3$[13].*

3 X-Ray Photon Correlation Spectroscopy

X-ray photon correlation spectroscopy, XPCS provides another method to measure the intermediate scattering function *S(Q,t)* on mesoscopic scales. Unlike NSE the accessed time domain for XPCS extends less towards the short time regime whereas correlation times of many seconds can be reached. XPCS is basically the same as photon correlation spectroscopy with visible light, also referred to as dynamic light scattering, DLS. It measures the intensity-intensity time auto-correlation function, g_2, of a speckle pattern in the scattered radiation, given by:

$$g_2\left(\vec{Q},t\right) = \left\langle I\left(\vec{Q},t_0\right) I\left(\vec{Q},t_0+t\right)\right\rangle / \left\langle I\left(\vec{Q},t_0\right)\right\rangle^2 = 1 + A(Q)\left[S\left(\vec{Q},t\right)/S\left(\vec{Q}\right)\right]^2 \qquad (21)$$

where t is the correlation (delay) time. For an ergodic system, i.e. a fluid but not a glass or gel, < …. > denotes likewise a time or ensemble average. 0 < A(Q) < 1 is an optical contrast factor that depends on the experimental setup details. For the second equality the Siegert relation has been used to relate $\left[S\left(\vec{Q},t\right)/S\left(\vec{Q}\right)\right]^2$ to the intensity autocorrelation function. The principle of the instrumental setup is relatively simple as shown in Fig.5. The visibility of the time dependent correlation function depends on the coherence properties of the X-ray beam, and the size of the illuminated sample volume. The size of the effective sample region determines the size of the speckles, i.e. the solid angle $\Delta\Omega$, which an intensity peak or minimum covers in

the random speckle texture of scattered intensity (see Fig.7). The detector element of area $\Delta x \Delta y$ must at most cover this solid angle, i. e. $\Delta\Omega \approx (\lambda/\Delta x)\,(\lambda/\Delta y)$. For example a sample of 30μm height (width) observed using a wavelength of λ=1.5Å and the detector at 5m distance allows for an opening height (width) of at most 25μm. If the detection area is enlarged the sample volume must be reduced. In other words a larger sample volume requires a smaller detection element. A way out of this intensity limiting trap is to use area detectors (CCD cameras) which have small-size individual pixels. However, these devices are not suited for online digital time correlation, but rather for time averaging as required to gain static information. Software correlation, on the other hand is limited to short times. There is clearly a need for fast solid-state detectors, like the Pilatus detector developed at the Paul-Scherrer Institute (PSI) in Switzerland. It has properties of a 2D detector, which allows for a reasonable number of pixel detection with each pixel calculating g_2 in real time with a time resolution of μs [14]. This offers new applications where an improved signal to noise ratio is required, such as the parallel variation of the Q-vector, the dynamics of crystallisation, to name just a few. Besides the detector pixel size, the ratio of source size and detector distance is another factor that enters the contrast A. The latter property is assigned to the beam as the transverse coherence length $\xi_{trans}=\xi_{x,y} = \lambda R/(2\pi\sigma_{x,y})$, which is to be compared to the sample size. The longitudinal coherence length of the beam is given by the degree of monochromatisation as $\Lambda_c = \lambda(\Delta E/E)/\pi$. For easily achievable monochromatisations of $\Delta E/E \approx 10^{-4}$, the effects of the transverse coherence lengths dominate up to $Q < 10^4\ 2\pi/\Delta x$. The required source characteristics, in combination with the necessary intensity are only provided by modern synchrotron radiation facilities. Indeed, from the mid nineties on such experiments were performed at all synchrotron machines of the third generation, where beam lines especially devoted for coherent experiments have been installed. In Europe these are the European Synchroton Radiation Facility (ESRF) in Grenoble/France, the Swiss Light Source (SLS) at PSI in Switzerland and in Hasylab at DESY in Germany.
In short, the coherence consideration boils down to the question on how the Q-resolution setup should be in order to resolve modulations of the size of the scattering volume (sample size).

The scattering contrast in the samples results from differences in the electron density which depend on the chemical structure. The very first successful experiments were therefore performed on colloidal metals [15]. However, as long as pure hydrocarbon particles without or with oxygen or nitrogen atoms included are considered, the scattering contrast is small and the intensities are correspondingly low. In these cases, area detection via a CCD camera is vital, but it limits the time resolution as discussed before. Using modern technologies with increased degree of coherence, beam-stability, number of (coherent) photons through the sample pinhole and higher energies etc., it is now possible to get to a sufficient signal-to-noise level even for soft matter samples, comparable to what is known from DLS.

Fig.5 shows the schematic layout of a XPCS beam line. A very thorough and up-to-date description of the method and very recent applications can be found in the review article by G. Grübel et al. in Ref.[16] and also at in the web-sites of the cSAXS beam line at the SLS in PSI in Switzerland and the Troika beam line ID10A at the ESRF in Grenoble in France [17].

XPCS is a new technique. Some inspection of recent publications shows how rapidly the field of application is growing. The topics span from biological imaging by hard X-ray diffraction microscopy [18], to the follow-up of atomic diffusion in $Cu_{90}Au_{10}$ by coherent X-

rays [19], the dynamics of lamellar phases studied by XPCS and dynamic light scattering [20], to the problem on how the elastic behaviour of a viscoelastic liquid can be investigated using grazing incidence XPCS [21]. Regarding the physics of colloidal systems, aging and slow dynamics have been investigated [22]. Last but not least the dynamics of charged colloidal particles has been studied [23].

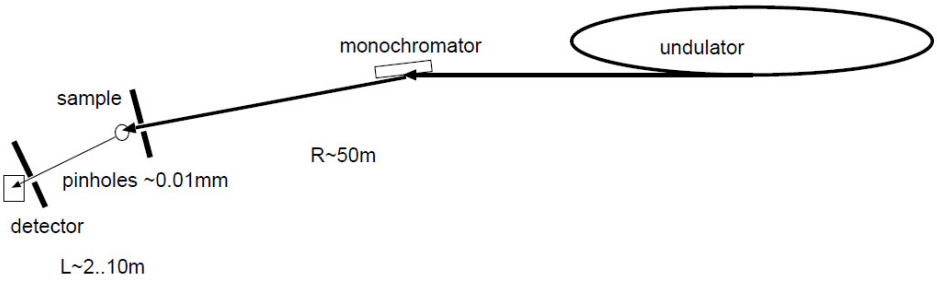


Fig. 5: *Schematic setup of a XPCS beam line at a synchrotron source. The effective X-ray source size is several 0.1mm. Besides a monochromator (e.g. a Si-single crystal), the instrument basically consists of a small pinhole in front of the sample defining the effective scattering volume plus a so-called guard slit which limits the primary beam so that only the small scattered intensity can be observed at the detector position. For a detailed discussion about the guard slit, see Fig.18-8 in Ref.16. A small pinhole in front of the detector limits the intensity collection to the solid angle to approximately the size of a typical speckle. For multispeckle detection with a CCD camera, the effective pinhole size is the size of one CCD pixel.*

It should be noted that the setup of a XPCS experiment is at the same time also an extremely good X-ray small-angle scattering machine. The regular SAXS spectrometers at synchrotron facilities are usually optimized for flux in order to achieve a sufficient short-time resolution for real-time experiments, for example for kinetic investigations or stopped flow experiments. At XPCS beam lines, the spatial resolution is unmatched since the contributions to the resolution, $\Delta Q/Q$, resulting from the finite slit width can be neglected. A typical example for the resolution capability is shown in Fig.8.

Another important issue is the radiation damage in polymeric, colloidal or biological samples caused by the X-rays. The local radiation dose applied to a small sample volume amounts to about 5-10kGy/s, for currently used beam lines (compared to less than 10^{-4} Gy/s for a typical neutron scattering experiment). In some cases, irradiation may be used to find interesting physics, as was observed for a charge-stabilized suspension of silica spheres in DMF with a concentration of about 160mg/g and in addition 20µM LiCl. This system shows a crystalline FCC six-fold symmetric structure, in agreement with the theoretical phase diagram prediction at the given number densities of particles and the salt concentration. This is shown in the upper left picture of Fig.6, where Bragg peaks are observed in the small-angle scattering configuration. Pictures were taken for 0.1s exposure time while the X-ray beam was on. After 1000s of exposure structure has melted (see the picture in upper right corner).

The structure recovers after the beam has been switched off, (see lower row in Fig.6) reaching the starting condition again after 2000s. Obviously, the X-ray beam has created screened type of charges within the sample, which lead to melting, in plausible agreement with theoretical calculations. Switching-off the beam leads to the recovery of the structure, since previously formed charges and/or radicals do recombine.

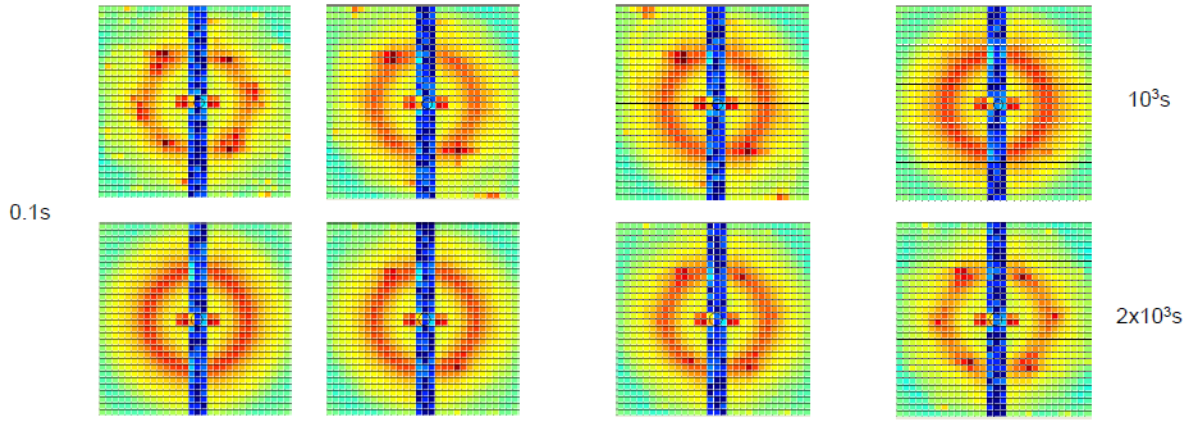


Fig. 6: *Structure melting and recovery from a Si/DMF colloidal suspension with 20μM LiCl added, under the influence of x-ray beam from cSAXS at SLS. The upper row shows melting occurring after 1000s, whereas the lower row shows recovery of the structure 2000s after the beam has been interrupted. Each picture was taken for the duration of 0.1s.*

3.1 Examples for X-ray Speckle Pattern

A speckle pattern was first investigated by Sutton and Mochrie et al. in their pioneering paper from 1991 [24] showing a Frauenhof-Diffraction pattern from a crystal, which clearly demonstrated the sufficient degree of coherence of the synchrotron X-ray beam. The authors predicted at the time of their discovery that dynamic experiments would become feasible.
As an example for a static scattering pattern, the X-ray speckle pattern from a porous silica glass is shown in Fig.7:

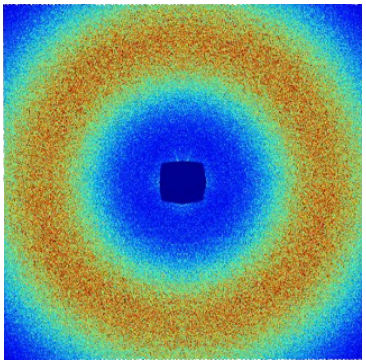


Fig. 7: *Static speckle pattern from a porous silica glass.*

The scattering of coherent light from a disordered system gives rise to a random diffraction or speckle pattern, long known from laser light scattering. The speckles are related to the exact spatial arrangement of the disorder. Such information is not accessible with incoherent light because the diffraction pattern observed in an ordinary diffraction experiment is typically an ensemble average containing only information on the average correlations in the sample. If the spatial arrangement changes with time, the corresponding speckle pattern also changes. The intensity fluctuations of the speckles thus provide information about the underlying dynamics (see section3.2.).

Consider again Fig.7. With a coherent beam of cross section ξ^2_{trans}, incident wave vector $\vec{k}$ and scattered wave vector $\vec{k}'$ scattered from a disordered sample. The instantaneous intensity *I* at a given point in the far field can be written as the square of a total field $\vec{E}(\vec{Q},t)$ according to

$$I(\vec{Q},t) = \left|\vec{E}(\vec{Q},t)\right|^2 = \left|\sum_n b_n(\vec{Q}) \exp\left[iQ \cdot \vec{r}_n(t)\right]\right|^2 \tag{22}$$

Here, $b_n(\vec{Q})$ is the scattering amplitude of the n-th scatterer located at position $\vec{r}_n$(t) at time t, and $\vec{Q} = \vec{k} - \vec{k}'$ in units of ħ. is the momentum transfer. The sum is taken over scatterers in the coherence volume, spanned by the transverse and longitudinal coherence lengths. The field in Eq.22 may be also expressed in terms of electron density function, ρ ($\vec{r}'$), as

$$\vec{E}(\vec{Q},t) = \int d\vec{r}' \, \rho(\vec{r}') \exp\left[i\vec{Q} \cdot \vec{r}'(t)\right] \tag{23}$$

A measurement of the intensity will naturally invoke a time average $< I(\vec{Q},t) >_T$ over the acquisition time T, but it does not involve any statistical ensemble average. If the system is not ergodic, i.e. has random static disorder, $< I(\vec{Q},t) >_T$ will display as a function of $\vec{Q}$ distinct and sharp variations in intensity, known as speckles. If, on the other hand, the system behaves ergodic, with fluctuations on time scales very short compared to the counting time, the measured time average is equivalent to an ensemble average, and $< I(\vec{Q},t) >_T$ can be replaced by the usual ensemble average denoted by $< I(\vec{Q},t) >$ The observed scattering is then featureless, apart from time averaged correlations in the sample similar to a regular scattering experiment with incoherent radiation. In Fig.7, a static speckle pattern from a porous silica glass is shown. The observed ring structure with a maximum in the intensity at $Q(I_{max})$ is due to pore-pore correlations in the glass. Static X-ray speckles have been observed in a large variety of different systems, i.e. porous systems [25], modulated bulk and surface structures [26], magnetic materials [27] and in systems exhibiting domain disorder [24].

3.2 Example for XPCS Dynamics

Long-time dynamics of charge-stabilized colloids: is there dynamic scaling or not?

Charge-stabilized suspensions are ubiquitously found in chemical and waste-treatment industries, and in medical and biological products. These systems are composed of meso-scaled colloidal particles dispersed in a low-molecular polar solvent like water. Examples include proteins and viruses, paint and clay particles, and well-characterized model systems consisting of spherical charge-stabilized particles. The beauty of the XPCS method is mainly due to the fact that different to dynamic light scattering, the sample can be completely opaque and may consist of particles with sizes in the order of the inverse Q-range accessible by the method. The study discussed now [28], takes advantage of this fact and deals with the question, on whether the dynamic scaling behaviour of the dynamic structure factor, observed experimentally for hard spheres in the work by Segrè and Pusey [29], holds true also for charge-stabilized particles. According to Segrè and Pusey, the decay of the dynamic structure factor of PMMA spheres for values Qd > 2.5, where d is the hard-sphere diameter, is controlled by self-diffusion according to:

$$S(Q,t)/S(Q) = \exp\left(-Q^2 W(t) \times D_s(Q)/D_s\right) \tag{24}$$

Here, W(t) is the particle mean-squared displacement (divided by 6) of short-time slope D_S, and $D_S(Q)$ characterizes the short-time exponential decay of S(Q,t). At long times, this factorization of the wave number and time dependence of S(Q,t) leads to a single-exponential decay, $S(Q,t)/S(Q) = \exp(- Q^2 D_L(Q) t)$, with a long-time diffusion function given by $D_L(Q) = D_S(Q) \times D_L/D_S$, where D_L is the long-time self-diffusion coefficient. It is especially interesting to perform XPCS experiments at a wave number $Q_s > Q_m$ where $S(Q_s) = 1$, and to check whether an estimate for the (long-time) self-diffusion coefficient can be obtained from this measurement as suggested by Pusey [30]. Note here that this kind of measurement could not been performed for this present system with dynamic light scattering, mainly for the reason that S(Q) can not be accessed within the Q-range of light scattering.

We have systematically explored the short-time and long-time dynamics of strongly interacting charged-stabilized particles in the fluid regime up to the liquid-crystal phase transition. The studied system consists of silica spheres of 164 nm diameter dispersed in dimethyl-formamide (DMF).

The X-ray synchrotron beam at energy of 8.7 keV from the cSAXS beam line at the SLS in PSI/Swiss was carefully collimated to reach high brilliance while keeping sufficient transversal coherence, a crucial requirement in the XPCS experiment. The sample was initially characterized by measuring the scattered intensity using the 2D Pilatus detector. Numerous visible minima in the high scattering vector *Q*-range provide a proof that the sample is very mono-disperse.

In further studies the point detector was used to measure the low-Q part of I(Q) required for calculation of the colloidal suspension structure factor *S*(Q). At selected Q-values, the time auto-correlation function was measured by means of a digital auto-correlator, see Fig.9. It

should be stressed that the quality of the correlation functions (in terms of the contrast and the noise) measured here by using XPCS correlation spectroscopy are comparable to those from using visible light.

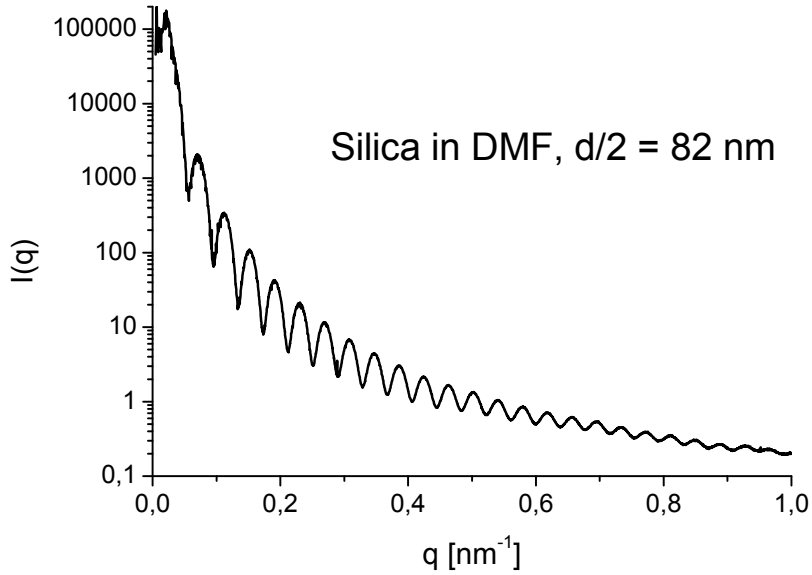


Fig. 8: *Pronounced multiple minima in the scattered intensity pattern prove that the sample is extremely mono-disperse.*

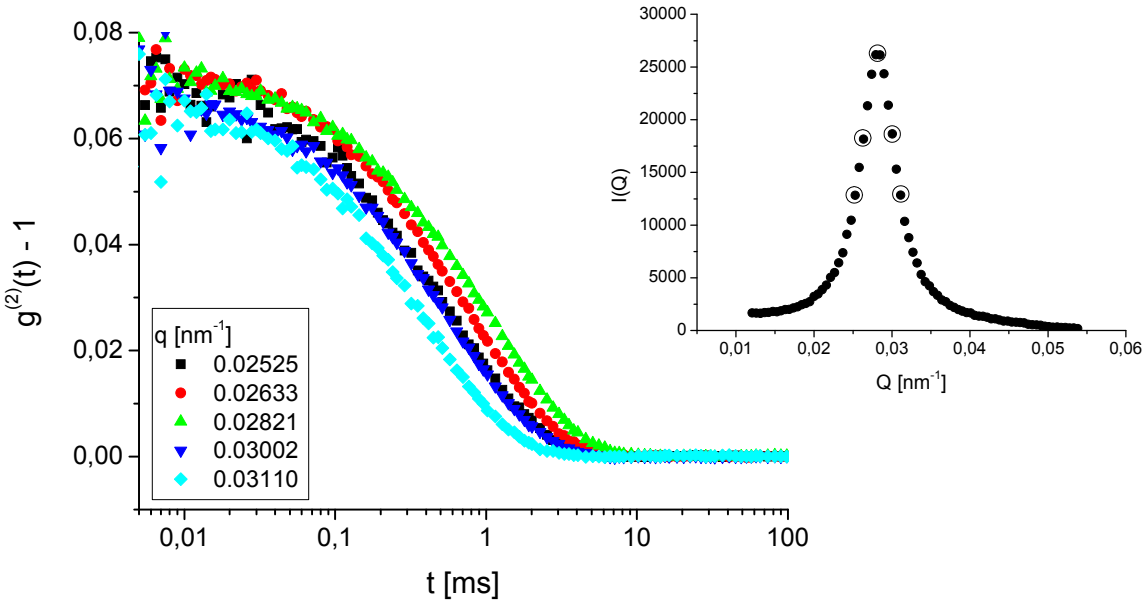


Fig. 9: *Intensity auto-correlation functions at Q-values as indicated. The circles in the inset are X-ray data for I(Q).*

Each correlation function was analyzed in terms of the initial decay of ln $g(t)$, as shown in Fig.10. The slope of ln $g_2(t)$ relates to $D(Q)Q^2$, where $D(Q)$ is the collective diffusion coefficient. From the initial decay part the short-time diffusion coefficient, $D_s(Q)$, was calculated, while from the slope of the long-time part, the long-time collective diffusion coefficient, $D_l(Q)$ was estimated, cf. Fig.10. The results of the fits are presented in Fig.11.
Following the scaling idea of Segre, Pusey and others [29], we have scaled the correlation functions according to the formula

$$f(Q,t) = \frac{\ln g^{(1)}(t)}{D_s(Q)\, Q^2} \tag{25}$$

In accordance with the scaling prediction, all functions measured at Q values not smaller than the peak position value follow a single master curve.

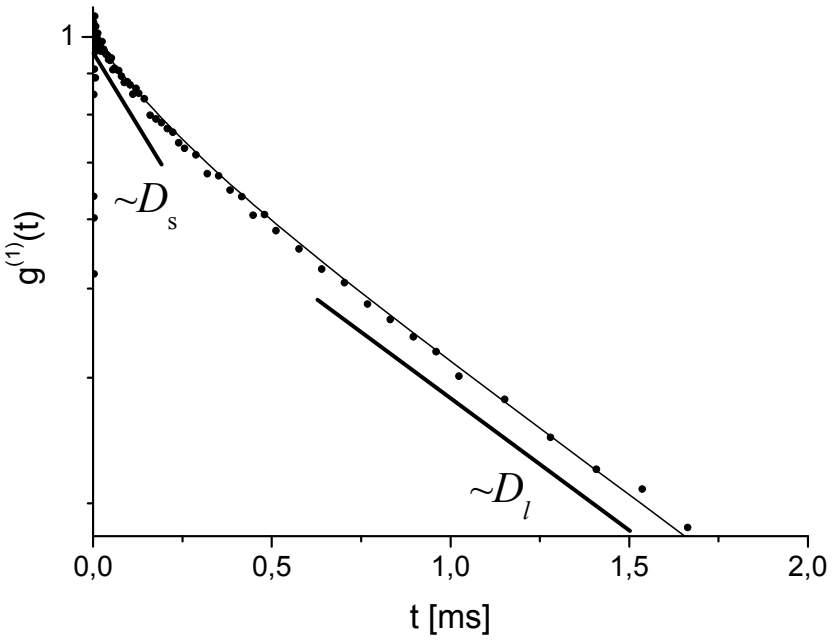


Fig. 10: *The initial decay of ln g(t). The thick lines indicate the slopes of the initial (short-time) and final (long-time) decay of ln g(t).*

We can conclude from our experimental findings that thanks to the extremely efficient XPCS setup at PSI, it has been possible to obtain high-quality time-correlation functions of a charged colloidal system. In our study, we have used a highly mono-disperse and strongly interacting system of silica spheres in DMF. The measurements of XPCS correlation functions for a series of chosen Q-values around the main structure peak position Q_m, yielded short- and long-time diffusion coefficients derived from the decay of ln $g(t)$. Scaling of ln $g(t)$

with $D_s(Q)$ Q^2 allowed for creating a master curve onto which all functions measured at Q-values not smaller than Q_m, collapse in agreement with the scaling prediction., cf. Figs.12a,b. The present XPCS result for the dynamic scaling behavior of the S(Q,t) for Silica in DMF nicely confirms corresponding DLS findings for a different system by Holmqvist and Nägele [31].

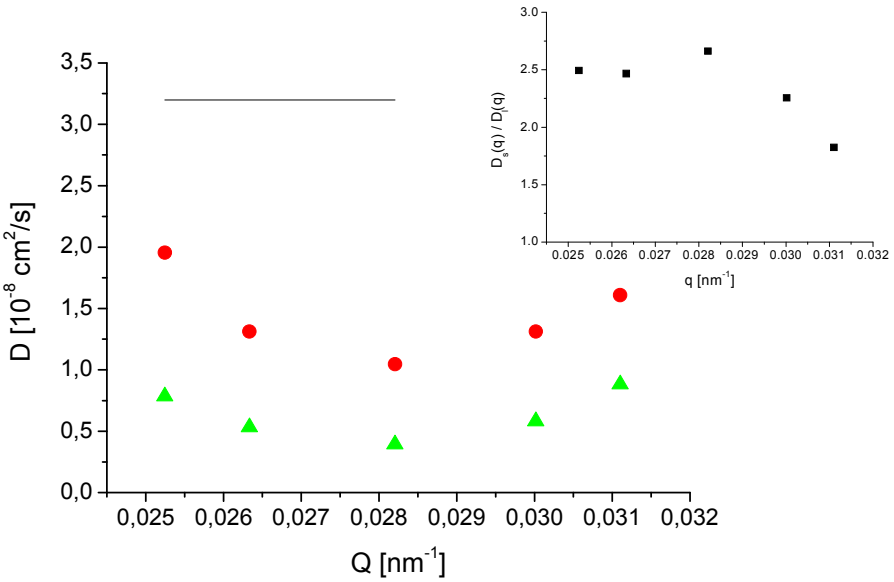
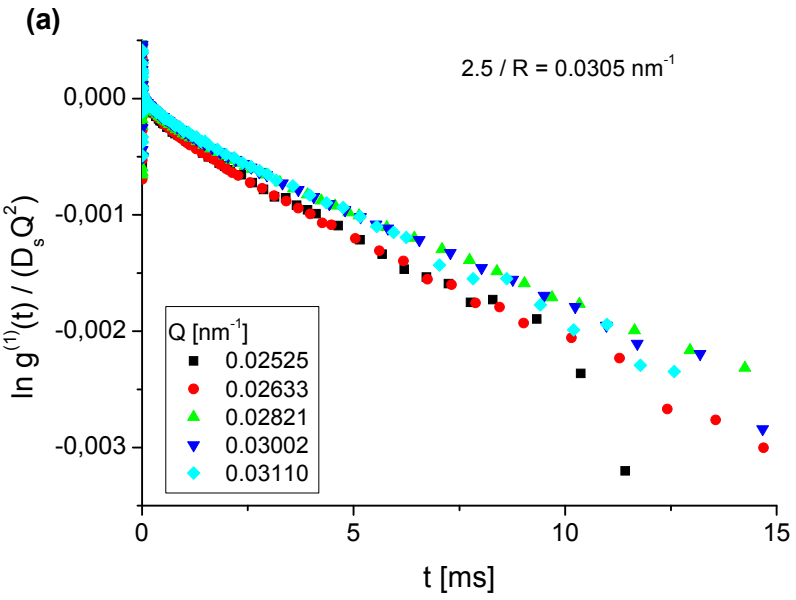


Fig. 11: *Values of short-time ($D_s(Q)$)(●) and long-time ($D_l(Q)$) (▲) collective diffusion coefficients, obtained from the slopes of the initial and final decay of ln g(t). (—) denotes the value for D_0. The inset shows the ratio of $D_l(Q)$ / $D_s(Q)$.*



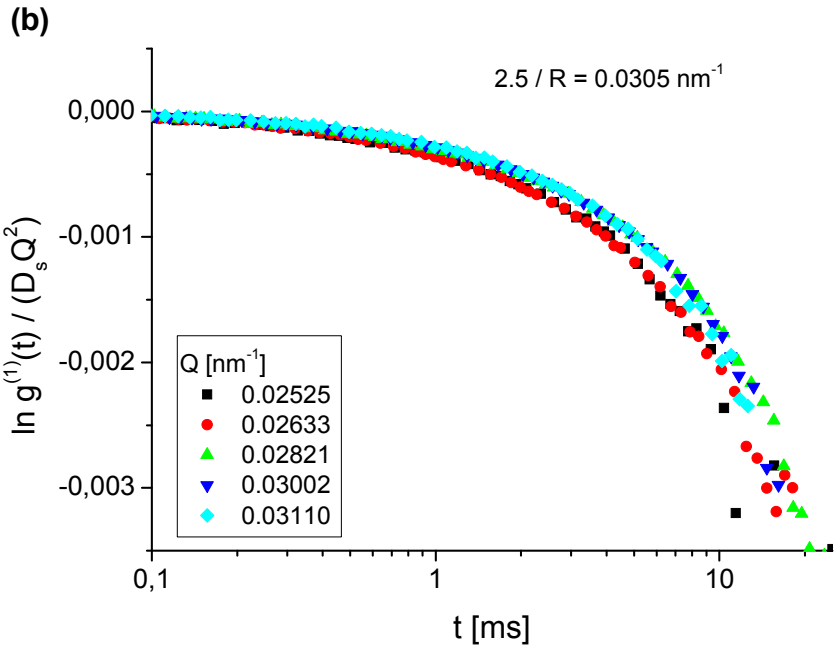


Fig. 12: *Scaled correlations functions measured at the Q-values of the peak of S(Q) and at Q-values lying symmetrically on both sides of the peak (cf. inset of Fig.9).*

References

[1] F. Mezei (ed.), Neutron Spin Echo, Lecture Notes in Physics, Vol. 128, Springer (1980)

[2] F. Mezei, C. Pappas and T. Gutberlet (eds.) Neutron Spin Echo Spectroscopy, Lecture Notes in Physics, Vol.601, Springer (2003)

[3] M. Monkenbusch, R. Schätzler and D. Richter, Nuclear Instruments & Methods In Physics Research Section A- Accelerators Spectrometers detectors And Associated Equipment **399**, 301 (1997)

[4] P. Rouse; J. Chem. Phys., **21**, 1250 (1953)

[5] M. Doi and S. Edwards, The Theory of Polymer Dynamics, Oxford University Press. Oxford (1994)

[6] M. Monkenbusch, IFF-springschool 2006, chapter D6

[7] M. Monkenbusch, O. Holderer, H. Frielinghaus, D. Byelov. J. Allgaier and D. Richter, J. Phys.: Condens. Matter **17**, 2903 (2005)

[8] G. Gompper and G. Schick in: *Phase Transitions and Critical Phenomena.* C. Domb and J. Lebowitz (eds.) Academic Press London (1994)

[9] M. Teubner and R. Strey, J. Chem. Phys. **87**, 3195 (1987)

[10] H. Endo, M. Mihailescu, M. Monkenbusch, J. Allgaier, D. Richter, B. Jakobs, T. Sottmann, R. Strey and I. Grillo, J. Chem. Phys. **115**, 580 (2001)

[11] A. Zilman and R. Granek, Phys. Rev. Lett. **77**, 4788 (1996)

[12] W. Helfrich, Z. Naturf. **28**, 693 (1973)

[13] O. Holderer, H. Frielinghaus, D. Byelov, M. Monkenbusch, J. Allgaier and D. Richter, J Chem. Phys. **122**, 094908 (2005)

[14] http://pilatus.web.psi.ch/pilatus.htm

[15] T. Thurn-Albrecht, W. Steffen, A. Patkowski, G. Meier, E. W. Fischer, G. Grübel and D. Abernathy, Phys. Rev. Lett. **77**, 5437 (1996)

[16] G. Grübel, A. Madsen and A. Robert in: *Soft Matter Characterisation.* R. Borsali and B. Pecora (eds.) Springer (2008) chapter 18, pages 954-995

[17] http://www.esrf.eu/UsersAndScience/Experiments/SoftMatter/ID10A/; http://www.psi.ch/sls/csaxs/csaxs

[18] E. Lima, L. Wiegart, P. Pernot, M. Howell, J. Timmins, F. Zontone and A. Madsen; Phys. Rev. Lett. **103**, 198102 (2009)

[19] M. Leitner, B. Sepiol, L.-M. Stadler, B. Pfau and G. Vogl, Nature Mat. **8**, 717 (2009)

[20] D. Constantin, G. Brotons, T. Salditt, É. Freyssingeas and A. Madsen, Phys Rev E **74**, 031706 (2006)

[21] Y. Chushkin, C. Caronna and A. Madsen, Europhys. Lett. **83**, 36001 (2008)

[22] A. Fluerasu, A. Moussaïd, A. Madsen and A. Schofield, Phys Rev.E **76**, 010401 (2007)

[23] A. Banchio, J. Gapinski, A. Patkowski, W. Häussler, A. Fluerasu, S. Sacanna, P. Holmqvist, G. Meier, M. P. Lettinga and G. Nägele, Phys. Rev. Lett. **96**, 138303 (2006)

[24] M. Sutton, S. G. J. Mochrie, T. Greytak, S. E. Nagler, L. E. Berman, G. A. Held and G. B. Stephenson, Nature **352**, 608 (1991)

[25] D. L. Abernathy, G. Grübel, S. Brauer, I. McNulty, G. B. Stephenson, S. G. J. Mochrie, A. R. Sandy, N. Mulders and M. Sutton, J. Synch. Rad. **5**, 37 (1998)

[26] Z. H. Cai, B. Lai, W. B. Yun, I. McNulty, K. G. Huang and T. P. Russel, Phys. Rev. Lett. **73**, 82 (1994)

[27] F. Yakhou, A. Letoublon, F. Livet, M. de Boissieu and F. Bley, J. Magn. Magn. Mater. **233**, 119 (2001)

[28] J. Gapinski, A. Patkowski and G. Meier, personal communication

[29] P.N. Segrè and P.N. Pusey, Phys. Rev. Lett. **77**, 771 (1996)

[30] P.N. Pusey, J. Phys. A: Math. Gen. **11**, 119 (1978)

[31] P. Holmqvist and G. Nägele, Phys. Rev. Lett. **104**, 058301 (2010)

D 9 Nuclear Resonance Scattering [1]

Rudolf Rüffer
European Synchrotron Radiation Facility
BP 220, F-38043 Grenoble, France

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Nuclear resonance scattering (NRS) with synchrotron radiation combines the outstanding properties of Mössbauer spectroscopy with those of synchrotron radiation. Since its first observation in 1984 [1] a rapid development of the technique and its applications followed which has just been reviewed [2, 3]. Thanks to the outstanding properties of 3rd generation synchrotron radiation sources NRS became nowadays an established spectroscopy on an atomistic scale which is element and even isotope sensitive and non-destructive. Applications comprise two main fields, *hyperfine spectroscopy* and *structural dynamics*. In hyperfine spectroscopy NRS is complementary to other nuclear techniques and yields useful information on atomic, magnetic, and electric structures. Those fields of applications benefit most which exploit the specific properties of synchrotron radiation: applications to high pressure, to grazing incidence geometry (surfaces and multilayers), to single crystals, and to very small samples. Structural dynamics on a ps to μs time scale as free or jump diffusion as well as rotational motions can directly be measured in the time domain by nuclear quasi-elastic scattering techniques. On the fast time scale the (partial) density of phonon states is directly accessible by nuclear inelastic scattering.

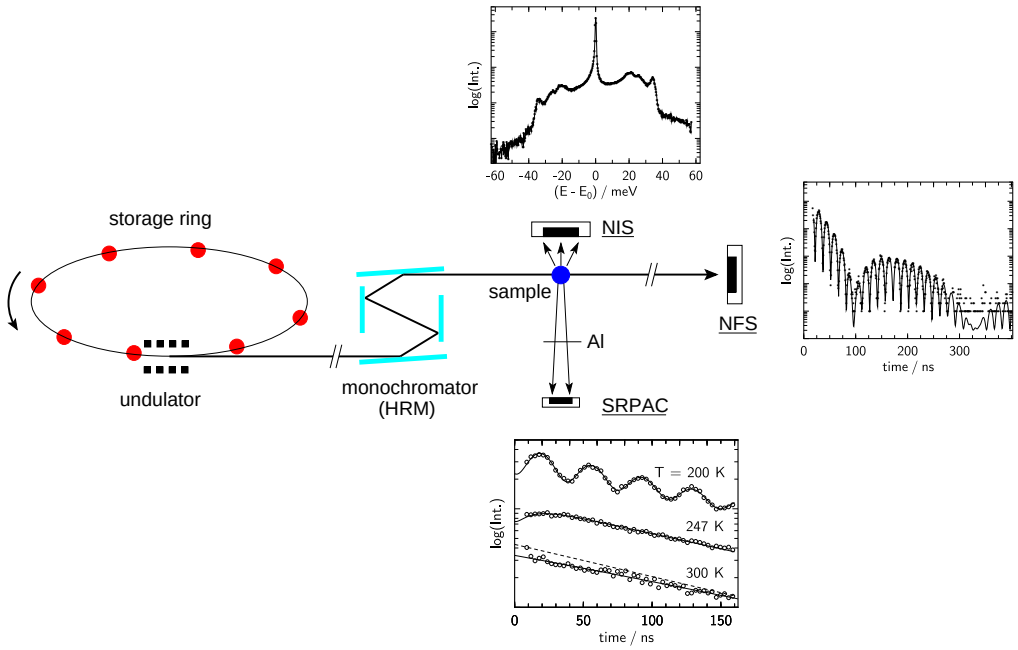


Fig. 1: *Scheme of the set-up for some NRS techniques: Nuclear Forward Scattering (NFS), Nuclear Inelastic Scattering (NIS), and Synchrotron Radiation based Perturbed Angular Correlation (SRPAC). Furthermore, the corresponding time and energy spectra are schematically shown measured by the detectors for NFS, SRPAC and NIS. The storage ring is operated in* few bunch *mode with e.g. 176 ns spacing between adjacent buckets (red bullets). The undulator produces the well collimated synchrotron radiation which is monochromized at the nuclear resonances by a high resolution monochromator, HRM (bandwidth $\Delta E \approx meV$). Finally the radiation impinges the sample. Depending on the scattering process and the scattering geometry different techniques are exploited. Variable sample environments allow for the combination of high/low temperature, high pressure, external magnetic field, and ultra high vacuum (UHV) conditions.*

2 Principles of Nuclear Resonance Scattering

Nuclear resonance scattering is a resonant x-ray scattering technique in the time domain. It is based on the Mössbauer effect and the scattering proceeds via the Mössbauer level whereas classical Mössbauer spectroscopy is an absorption spectroscopy and is normally carried out in the energy domain. Both are, in principle, connected by the Heisenberg uncertainty principle.
In case of Mössbauer spectroscopy a radioactive source provides the γ-quanta which are very sharp in energy (neV - μeV) and by varying the energy one measures absorption spectra behind the sample. In nuclear resonance scattering an x-ray pulse from a synchrotron radiation source which is very sharp in time (~ 100 ps) and very broad in energy (white radiation) is exciting the nuclear levels in the sample. The successive decay of these levels gives rise to an exponential intensity decay in time. This decay corresponds to the Lorentzian line in Mössbauer absorption spectroscopy.
There are several techniques utilizing nuclear resonance scattering: nuclear forward scattering (NFS), nuclear quasi-elastic scattering (NQES), nuclear Bragg diffraction (NBD), nuclear reflectometry (NR, for the investigation of surfaces and multilayers), nuclear small angle scattering (NSAS, measuring e.g. magnetic domain structures), nuclear inelastic scattering (NIS), and synchrotron radiation based perturbed angular correlation (SRPAC). In the following a short introduction will be given to the main techniques: nuclear forward scattering, nuclear quasi-elastic scattering, and nuclear inelastic scattering (see Fig. 1).

2.1 Nuclear forward scattering

Nuclear forward scattering [4, 5] (and nuclear reflectometry) has the closest analogy to Mössbauer spectroscopy in fact it is its scattering variant. The main domain is the determination of hyperfine interaction parameters such as internal magnetic fields (magnetic hyperfine field), electric field gradients, isomer shifts and the Lamb-Mössbauer factor (f_{LM}).
The set-up for NFS (see Fig. 1) looks also very similar to a standard set-up for Mössbauer spectroscopy. However, while Mössbauer spectroscopy is an absorption spectroscopy (one γ-quantum is absorbed by one nucleus) NFS is a scattering spectroscopy. The 'white' synchrotron radiation excites all Mössbauer levels in the sample and creates a *coherent collective nuclear state*. In the static case this nuclear state will decay in the forward direction giving rise to an excess of intensity at delayed times (see Fig. 1 NFS). The time scale is determined by the lifetime τ_0 of the involved nuclear level. Multiple scattering may influence the measured time response giving rise to *dynamical beats*. Furthermore, in case of split nuclear levels due to hyperfine interaction (electric, magnetic) an additional interference pattern, the so-called *quantum beat* structure is superimposed.

Dynamical beats: For samples with a large effective thickness, $t_{eff} = n_{MB} \cdot d \cdot \sigma_0 \cdot f_{LM}$, with n_{MB} the density of Mössbauer nuclei, d the geometrical thickness, σ_0 the resonant Mössbauer cross section, and f_{LM} the Lamb-Mössbauer factor, a *speed-up* effect is observed and dynamical beats (*Bessel beats*) show up in the time spectra.
In Fig. 2 the situation is shown for Mössbauer and NFS spectra. For a thin sample, $t_{eff} = 1$, we have a single Lorentzian line in the Mössbauer spectrum which corresponds to an exponential decay in the NFS spectrum (Fig. 2 a,b, dotted lines). Increasing the effective thickness ($t_{eff} = 25$) the Mössbauer line becomes much wider and non-Lorentzian (Fig. 2 a, solid line). In the NFS spectrum (Fig. 2 b, solid line) we observe two features: (1) a speed-up, showing up as an increase in intensity at early times (from 10^7 to 10^9 units in the present example) and (2) a

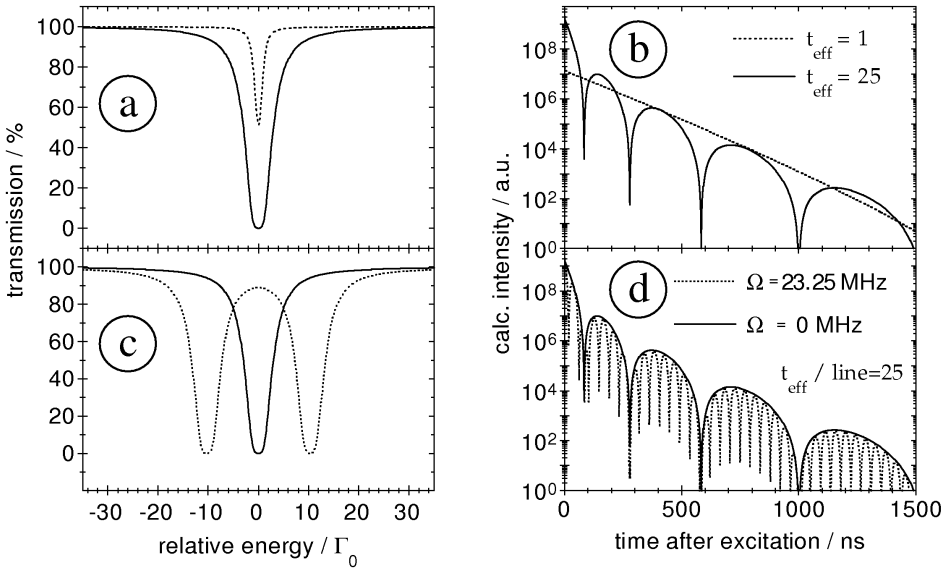


Fig. 2: *Simulations of corresponding spectra of Mössbauer spectroscopy (left panel, with* $\Gamma_0=\hbar/\tau_0$*) and NFS (right panel) for thin (*t_{eff} *= 1) and thick (*t_{eff} *= 25) samples. For the single lines (panel a and b) the thin sample leads to an exponential decay whereas the spectrum from the thick sample is further modulated by the dynamical beats. In panel c the energy spectra of a thick sample (*t_{eff} *= 25) with one (solid line) and two (dotted line) transition lines, respectively, are shown. In panel d the corresponding time spectra are shown. The sample with the two lines shows in addition to the dynamical beats the fast, equidistant quantum beat structure superimposed (from [6]).*

dynamical beat structure is superimposed to the decay. From this beat structure the effective thickness and correspondingly $\mathrm{f_{LM}}$ can very precisely be determined.

Quantum beats: As is well known hyperfine interaction might split the nuclear levels. In Mössbauer spectroscopy several absorption lines reveal this splitting whereas in NRS an interference pattern, the quantum beat structure, shows up.
In case of quadrupolar interaction the excited nuclear state (e.g. in case of ^{57}Fe) splits to the $\pm$3/2 and $\pm$1/2 levels. This gives rise to two absorption lines in Mössbauer spectroscopy (see dotted line in Fig. 2 c). The corresponding spectrum in NFS shows up as an interference pattern of these two transitions with a single frequency Ω of the quantum beats (see dotted line in Fig. 2 d). Due to the thick sample with an effective thickness $\mathrm{t_{eff}} = 25$, the dynamical beat structure is strongly modulating the quantum beats as an envelope. It is clearly seen that the quantum beat structure is equidistantly spaced whereas for the dynamical beat structure the distance of the minima increases with time. This can be described in a good approximation by

$$I_{NFS}(t) \propto \frac{t_{\mathrm{eff}}}{t/\tau_0} \cdot \cos^2\left(\frac{\Omega t}{2}\right) \cdot e^{-\frac{t}{\tau_0}} \cdot J_1\left(\sqrt{4t_{\mathrm{eff}} \cdot t/\tau_0}\right)^2 . \qquad (1)$$

The term $\cos^2(\frac{1}{2}\Omega \cdot t)$ describes the quantum beats, the Bessel function $\mathrm{J_1}$ the dynamical beats, and the exponential accounts for the overall exponential decay. For comparison the unsplit case is shown as solid lines in the same figures. Changing the strength of the hyperfine interaction will result in a different splitting and correspondingly in a different quantum beat frequency Ω. In case of magnetic hyperfine interaction full splitting of the nuclear levels occurs giving rise to

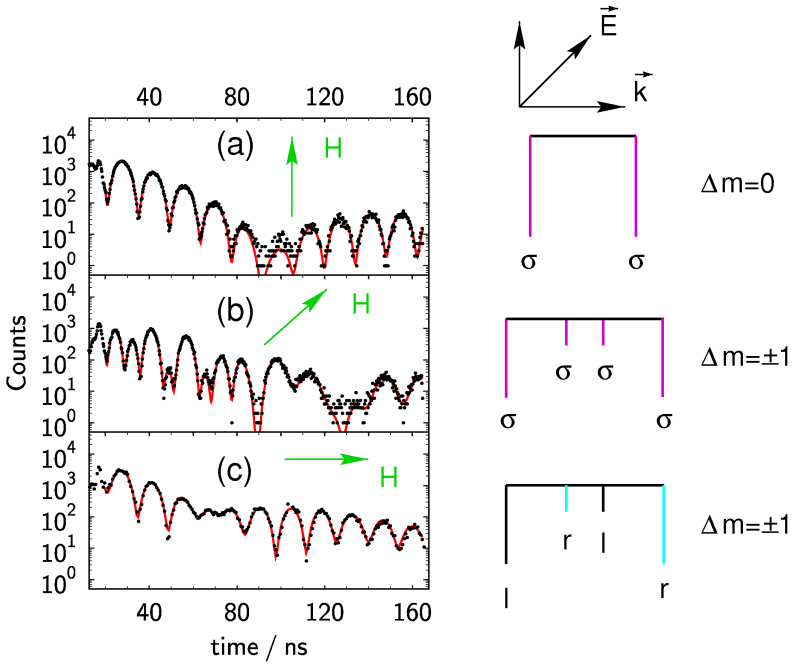


Fig. 3: *left panel: Measured time spectra of iron (NFS) in case of magnetic hyperfine interaction for various alignments of the hyperfine field $\vec{H}$ with respect to the wave vector $\vec{k}$ and the polarization $\vec{E}$. Solid lines are fits according to the full theory. right panel: nuclear transition lines with their polarization state, σ - σ linearly polarized, l and r - left and right hand circular polarized, respectively. Δm - change of magnetic quantum number (from [7]).*

six nuclear transitions and correspondingly to six absorption lines in Mössbauer spectroscopy with ^{57}Fe. In NRS spectroscopy a more detailed interference pattern will result. Contrary to Mössbauer spectroscopy where the γ-rays from the radioactive source are normally unpolarized now the x-rays from the synchrotron radiation source are highly linearly polarized. This feature strongly modifies the time spectra which are displayed in Fig. 3. The important parameters are the orientation of the three vectors with respect to each other, the wave vector $\vec{k}$, the polarization vector $\vec{E}$, and the hyperfine field vector $\vec{H}$.
If all three vectors are perpendicular to each other a simple quantum beat pattern with one single frequency and high contrast results originating from the two equally strong $\Delta m = 0$ transitions (Fig. 3 a). A similar spectrum with only one frequency, however, with less contrast appears when $\vec{H} \parallel \vec{k}$. In this case the two inequally strong $\Delta m = +1$ and the two inequally strong $\Delta m = -1$ transitions interfere independently giving now rise to left and right hand circular polarization (Fig. 3 c). Finally, for $\vec{H} \parallel \vec{E}$ all $\Delta m = \pm 1$ transitions interfere giving rise to a more complicated spectrum which is sigma polarized (Fig. 3 b). The slow overall modulation is caused by dynamical beats due to the finite effective thickness.

2.2 Nuclear quasi-elastic scattering

Nuclear quasi-elastic scattering measures structural dynamics on a ps to μs time scale. The coherent and the incoherent channel can be utilized. In the first case, the coherent channel, the Lamb-Mössbauer factor has to be greater than zero ($f_{LM} > 0$). The set-up is the same as in NFS (see Fig. 1). The incoming x-ray pulse creates a coherent collective nuclear state

which decays in the static case in forward direction. Dynamics, e.g. the jump of a Mössbauer nucleus from one atomic site to another (in space and angle see e.g. [8]), destroys this state. As a consequence no x-ray is scattered in forward direction, i.e., the measured intensity in the NFS detector is decreased at later times. We will get an 'accelerated decay' or a 'damping' of the NFS intensity I_{NFS}(t) which might be described in a simplified picture by:

$$I(t) \propto I_{NFS}(t) \cdot e^{-2\lambda_t t} \cdot e^{-\lambda_r t} \tag{2}$$

The first exponential is related to the van Hove self-intermediate function with λ_t being the translational relaxation rate and the second one to the rotational correlation function with λ_r being the rotational relaxation rate.
In the second case, the incoherent channel, the scattering is independent of the Lamb-Mössbauer factor, f_{LM}. The set-up (see Fig. 1) is similar as in traditional time differential perturbed angular correlation spectroscopy. Therefore the technique is called synchrotron radiation based perturbed angular correlation, SRPAC. The incoming x-ray pulse selectively excites a single nucleus which decays in the static case with an angular distribution according to the anisotropy parameter, A_{22}. Dynamics, e.g. rotational motion monitored by the electric hyperfine interaction $\hbar\Omega$, changes this distribution and gives rise to a damping of the intensity signal. The scattering intensity can be written as

$$I(t) = I_0 \cdot e^{-t/\tau_0}\{1 - A_{22}G_{22}(t)\} \tag{3}$$

with A_{22} the anisotropy and angular term and G_{22} the perturbation factor. It reduces in the slow approximation to

$$G_{22}(t) \propto e^{-\lambda_r t} \cdot \cos^2\left(\frac{\Omega t}{2}\right) \tag{4}$$

and in the fast approximation to

$$G_{22}(t) \propto e^{-(\Omega^2/\lambda_r)t}. \tag{5}$$

Combining both techniques the translational and rotational relaxation rates can be separately extracted.

2.3 Nuclear inelastic scattering

Nuclear inelastic scattering directly measures the *(partial) density of phonon states* [9, 10, 11]. The principle set-up is shown in Fig. 1 (NIS). The synchrotron radiation is monochromized by a scanning high resolution monochromator (HRM) with constant energy resolution in the (sub-) meV (1 meV $\hat{=}$ 0.24 THz) range and at the same time covering a large energy range of several hundreds of meV. At resonance the x-ray pulse creates a coherent collective nuclear state as in NFS which decays either in forward direction or incoherently, due to internal conversion, spin flop etc., in the entire solid angle of 4π. This feature gives a simple and effective method at hand to measure the instrumental function (detector NFS) in parallel with the inelastic spectrum (detector NIS). While scanning the HRM the nuclear resonance can only be excited when at the same time a phonon is excited or annihilated. In NIS this is a purely incoherent process with the consequence of a perfect averaging over the momentum $\vec{q}$. The successive nuclear decay proceeds as a single exponential in the entire solid angle covered by the '4π'-detector (detector NIS). In practice only a solid angle of about π can be achieved.
As an example the energy dependence of nuclear inelastic scattering of synchrotron radiation in a polycrystalline α-iron sample at room temperature is shown as a function of energy of the

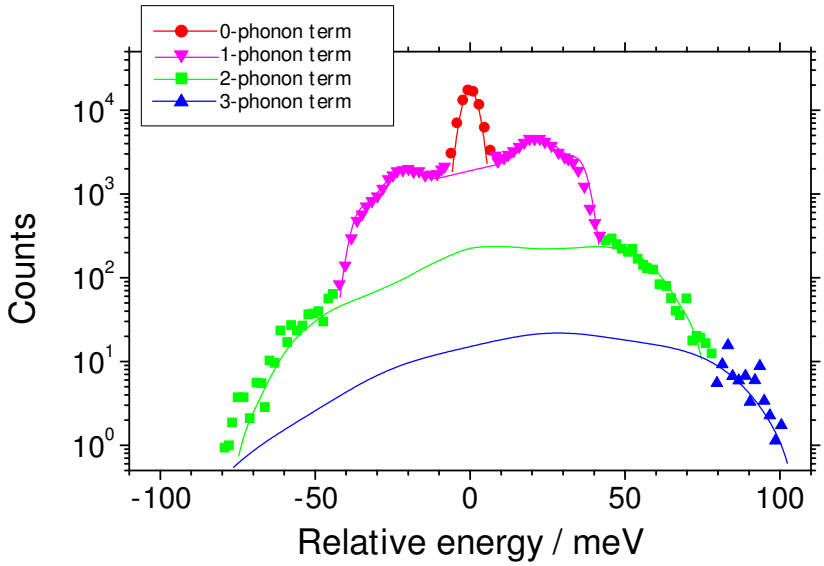


Fig. 4: *Expansion of the energy spectrum of nuclear inelastic scattering of synchrotron radiation in α-iron in multi-phonon terms. The data were taken at room temperature. Different symbols show the regions of the spectra, where the corresponding contributions are dominant. The lines are the calculations according to Eqs.6-8 and convoluted with the instrumental function of the monochromator (from [12]).*

incident radiation (Fig. 4). The central peak corresponds to elastic scattering. The structure beyond the central peak shows the energy spectrum of inelastic scattering, accompanied either by creation ($E > 0$) or by annihilation ($E < 0$) of phonons. At ambient temperature one may recognize various contributions to the energy spectrum, which correspond to inelastic scattering accompanied by excitation or annihilation of different number of phonons. The normalized probability of nuclear inelastic scattering $W(E)$ can be decomposed in terms of a multiphonon expansion [13]

$$W(E) = f_{\mathrm{LM}} \left(\delta(E) + \sum_{n=1}^{\infty} S_n(E) \right). \tag{6}$$

The Dirac δ-function $\delta(E)$ describes the elastic part of scattering (zero-phonon term), and the n-th term of the series $S_n(E)$ represents the inelastic scattering accompanied by creation (annihilation) of n phonons. The one-phonon term is given by

$$S_1(E) = \frac{E_R \cdot g(|E|)}{E \cdot (1 - \mathrm{e}^{-\beta E})}, \tag{7}$$

and the subsequent terms under harmonic approximation may be found through the recursive relation:

$$S_n(E) = \frac{1}{n} \int_{-\infty}^{\infty} S_1(E') \, S_{n-1}(E - E') \, \mathrm{d}E'. \tag{8}$$

Here $\beta = (k_{\mathrm{B}} T)^{-1}$ with k_{B} the Boltzmann constant, T the temperature; $E_{\mathrm{R}} = \hbar^2 k^2 / 2M$ the recoil energy of a free nucleus; k the wave vector of the x-ray quantum; M the mass of the

atom. The function $g(E)$ is the normalized density of phonon states

$$g(E) = V_0 \frac{1}{(2\pi)^3} \sum_j \int d\vec{q}\, \delta\left[E - \hbar\omega_j(\vec{q})\right], \tag{9}$$

where V_0 is the volume of unit cell, index j numerates the branches of the dispersion relation $\hbar\omega_j(\vec{q})$, $\vec{q}$ is the phonon momentum, and the integral is taken within the first Brillouin zone. The detailed theory of nuclear inelastic scattering has been published by several authors [13, 14, 15]. From the density of phonon states other (thermo)dynamic quantities can be derived as the vibrational contribution to the internal energy, lattice specific heat at constant volume and pressure, vibrational entropy. From the sum rules [16] the density of phonon states, f_{LM} and the mean square displacement, the mean kinetic energy, the mean force constant, and the second order Doppler shift can be determined.
In summary, NIS gives direct access to the density of phonon states and various (thermo)dynamic quantities. It is complementary to methods as inelastic neutron, x-ray, and light scattering. In those techniques mainly dispersion relations are measured which are fitted to a model and in a final step the density of phonon states can be derived.

3 Applications to high pressure

High pressure applications are one of the domains which benefit most from the outstanding properties of synchrotron radiation. The small beam size allows one to work with samples as small as 25 μm in diameter. In that case state-of-the-art diamond anvil cells (DAC) are able to reach pressures far above the 100 GPa regime up to 300 GPa, those pressures which are common in the center of the Earth. In combination with laser heating (5000 K) access to the entire relevant (pressure, temperature)-space becomes available. First applications have been reported for iron containing specimens [17, 18]. Furthermore, the small divergence of the beam allows sophisticated sample environments, in situ pressure calibration and diffraction studies not only for the determination of the induced volume change with pressure but also to identify phase transitions. The following two examples demonstrate typical applications. In the first example iron metal has been investigated by NFS and NIS in order to follow over the pressure induced phase transition magnetic order (via the hyperfine field) and structural dynamics (via the (partial) density of vibrational states, DOS) and in the second example the long standing quest for the magnetic phase in SmS under applied pressure has been tackled.

3.1 Iron metal

Recently, the high pressure behaviour of iron gained much interest due to unusual superconductivity, to simulations of the inner core of the Earth, and to improved technical developments in high pressure research. It is well known that α-iron undergoes under pressure a magnetic and structural phase transition to the hexagonal closed-packed (hcp) phase of iron (ϵ-iron) in the pressure range of 10 to 22 GPa at ambient temperature.
A 2.5 μm thick iron foil enriched to 95% in ^{57}Fe was pressurized in a DAC [19]. Figure 5 left panel shows the NFS spectra for various pressures. At 3 GPa the spectrum reveals iron in the pure α-phase witnessed by the regular quantum beat pattern from the internal hyperfine field of 33 T. At 14 GPa the influence of the ϵ-phase is already quite pronounced. With a further increase in pressure the influence of the non-magnetic ϵ-phase becomes dominant. At 21 GPa the sample

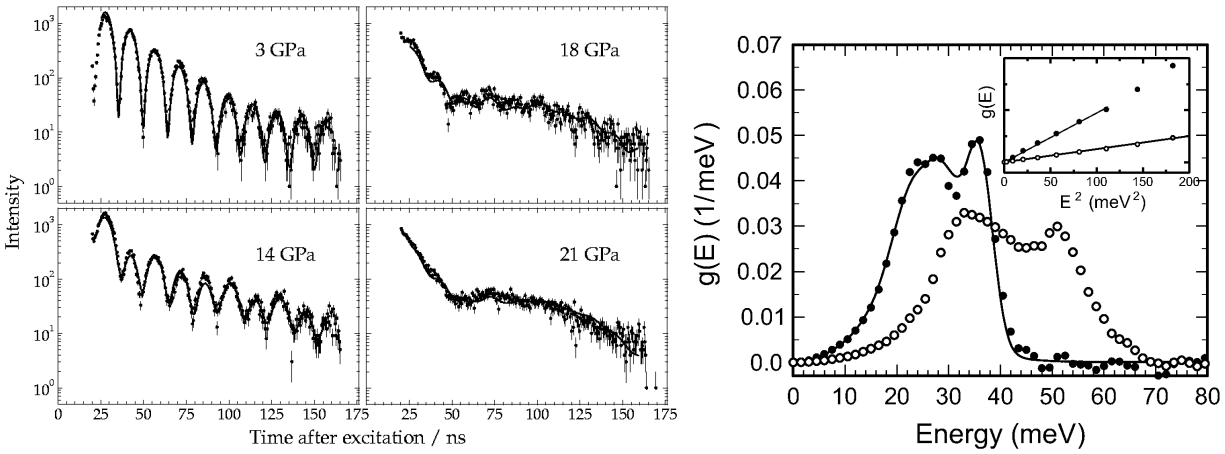


Fig. 5: *The pressure induced $\alpha \longrightarrow \epsilon$ transition in Fe. left panel: Time spectra of NFS at various pressures (from [19]). The solid line represent the fit to the data. right panel: Phonon density of states at ambient pressure (solid circles) and 42 GPa (open circles). The solid line represents the DOS from neutron scattering. The inset displays the corresponding low energy part for the determination of sound velocities (from [20]). Here the solid lines represent a linear fit to the data.*

is almost completely in the ϵ-phase. Now the NFS spectrum is modified only by a single Bessel minimum at 50 ns which originates from the increased effective thickness in comparison to that of the magnetically split one.
With the same sample, NIS energy dependencies have been recorded and the vibrational densities of states have been extracted [20]. Examples at ambient pressure and at 42 GPa are displayed in Fig. 5 right panel. All spectral features of the DOS of ϵ-Fe are shifted to higher energies with respect to α-Fe revealing a hardening of the lattice vibrations. Several thermodynamical properties have been derived as the Lamb-Mössbauer factor, f_{LM} and the mean-square displacement, respectively, the mean phonon energy and the corresponding Debye temperature, the lattice contribution to the specific heat c_{V}, the internal energy, entropy and the mean force constant. Finally, from the low-energy part of the density of phonon states the average velocity of sound v_{av} (Fig. 5 right panel inset) has been derived according to the relation

$$g(E) = \frac{V}{2\pi^2\,\hbar^3\,v_{\mathrm{av}}^3}\,E^2 \tag{10}$$

with *V* the volume per Fe atom. Extensions to higher pressures have been reported [21] and investigations to account for anisotropies [22, 23]. These sound velocity data have been complemented by recent data from inelastic x-ray scattering [24] and in the very high pressure regime by those from shock waves [25].

3.2 Magnetic Order in SmS

Recently, magnetic quantum critical points (QCPs) have attracted much interest, especially after the discovery of the appearance of unconventional superconductivity at a QCP and of the causative role of spin and charge fluctuations.
One class of systems known as Kondo insulators deserves special attention. Among them, SmS is known to be a non-magnetic (Sm^{2+}) semiconductor at ambient pressure (black phase). It

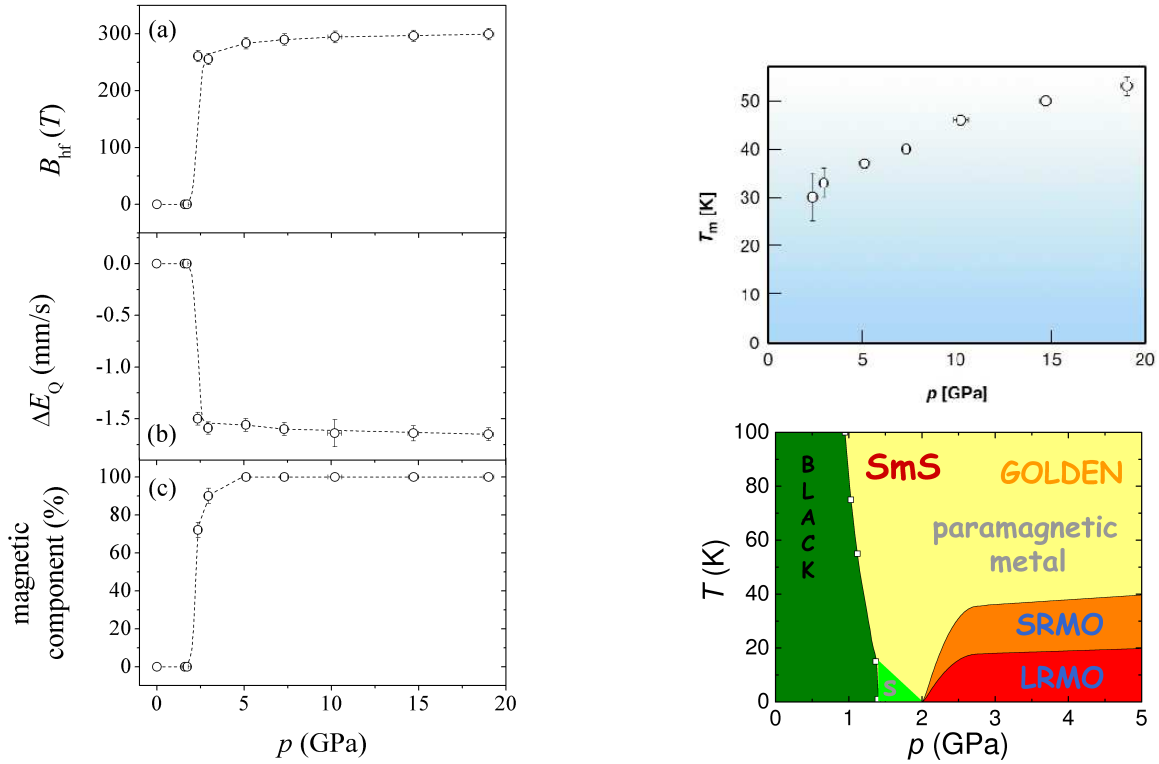


Fig. 6: *left panel: Pressure dependence for SmS of (a) the magnetic hyperfine field (B_{hf}), (b) the electric quadrupole splitting (ΔE_Q), and (c) the magnetic component fraction at T = 3 K. The dashed lines through the data points are guides to the eye.*
right panel top: Pressure dependence of the ordering temperature T_m. right panel bottom: The phase diagrams of SmS as a function of temperature and pressure, with the magnetically-ordered phases as discovered by NFS. Legend: SRMO = short-range magnetic order, LRMO = long-range magnetic order, S = semiconductor. (figures taken from [26, 27])

becomes metallic (golden phase) under pressure (0.65 GPa at room temperature), at which the Sm ions are in an intermediate valence state. So far it has been considered that, by increasing further the pressure, the golden phase could evolve into a magnetic state where the Sm ions are trivalent (Sm^{3+}). However, to date there has been no clear evidence of it.
Figure 6 (left panel) shows the pressure dependence of the hyperfine parameters (magnetic hyperfine field B_{hf} and electric quadrupole splitting ΔE_Q) and of the magnetic component fraction at 3 K [28]. The pressure-induced phase transition from a non-magnetic state into a magnetically-ordered state at $p_c \sim 2$ GPa is evident. The steep variation of both B_{hf} and ΔE_Q as well as the coexistence of the low and high-pressure phases in the vicinity of p_c point towards a first-order transition. The absolute values of B_{hf} and ΔE_Q are consistent with the occurrence of a Γ_8 crystal field ground state for the Sm^{3+} ions, with a magnetic moment of the order of 0.5 μ_B. The high-pressure magnetically-ordered phase is stable up to at least 19 GPa, as shown by the very weak pressure dependence of the hyperfine parameters and by the increase with pressure of the ordering temperature T_m (see Fig. 6 right panel top). These results give the first direct evidence that a magnetic ground state appears for SmS at a pressure of 2 GPa. Surprisingly, at this pressure the valence is about 2.8 and reaches the trivalent state only at about 13 GPa [29] without any visible change in the magnetic properties. Figure 6 (right panel bottom) displays the proposed phase diagram for SmS.

4 Applications to nano-scale materials

4.1 Dynamics of nano-crystalline material

Fast and slow dynamics, i.e., vibrational dynamics and diffusion, in amorphous, low-dimensional and nanostructured solids is a rapidly growing research field. Atomic diffusion governs the structural relaxation and associated changes in the physical properties whereas vibrational dynamics becomes crucial in thermodynamic, conductive, and optical properties of mesoscopic systems.

Scale Effects on the Vibrational Properties of Nanocrystalline Iron: A wealth of novel phenomena, such as phonon folding, phonon confinement, and vibrational interface states have been revealed in the last few years. A key role in this respect is played by the density of vibrational states, g(E). In the DOS of nanocrystalline (NC) materials, i.e., polycrystals with an ultrafine grain size d, both finite-size effects and interface-related phenomena due to the large fraction of disordered interfaces are expected. In particular, two basic questions are extremely relevant: (i) how does the DOS scale with d, and (ii) is there a deviation from the usual Debye law $g(E) = a \cdot E^2$ at low energies?

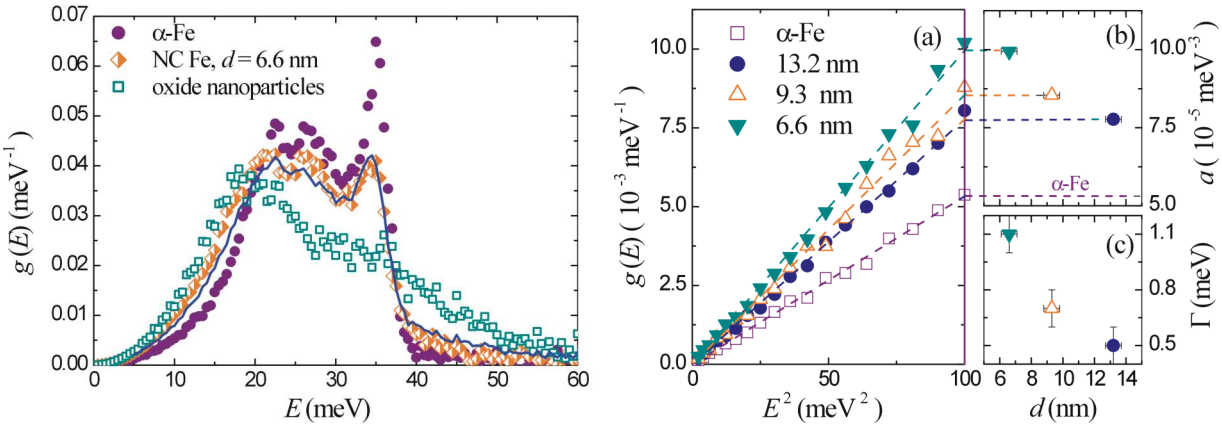


Fig. 7: *(left panel) DOS of NC iron (d = 6.6 nm), iron oxide nanoparticles, and reference α-iron. Solid line: linear combination* $x \cdot g_{\alpha-Fe} + (1\text{-}x) \cdot g_{oxide}$ *after convoluting* $g_{\alpha-Fe}$ *with a 1.1 meV FWHM Lorentzian in order to represent phonon damping). (right panel) (a) low-energy DOS vs. energy square of NC samples with different grain size. Dashed lines: best fits according to* $g(E) = a \cdot E^2$*; (b)-(c) coefficient a and damping parameter* Γ *vs. grain size, respectively (figures taken from [30]).*

Nuclear inelastic scattering of synchrotron radiation was applied to the determination of the DOS in NC iron samples with different grain size, prepared by inert gas condensation [31]. The oxidation degree of the samples (1-x) is defined as the atomic fraction of oxidized Fe atoms, obtained from X-ray diffraction and Mössbauer spectroscopy. In the NC samples, in comparison with a reference α-Fe foil, an enhanced population of low-energy modes and a broadening of the longitudinal peak at about 36 meV were observed (Figure 7 left panel). This softening can be explained only partially by oxidation of the grain boundaries: a linear combination $x \cdot g_{\alpha-Fe} + (1\text{-}x) \cdot g_{oxide}$ does not fully account for the low-energy intensity. Here $g_{\alpha-Fe}$ and g_{oxide} is the DOS measured for α-Fe and totally oxidized nanoparticles, respectively. The remaining excess modes are most likely ascribable to vibrations of atoms at the grain boundaries, with modified local environment and softened force constants. In a very recent work [32] this could

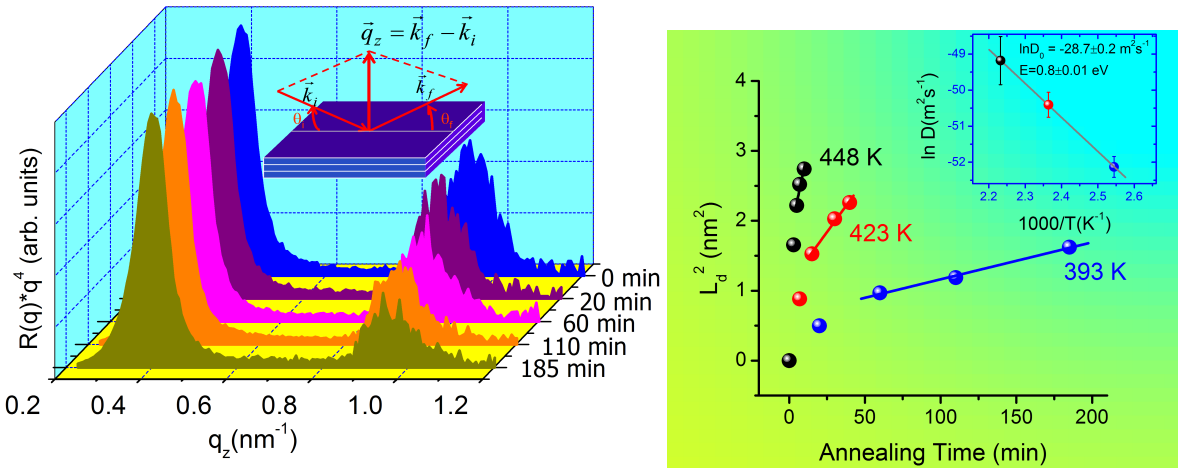


Fig. 8: *(left panel) Nuclear resonance reflectivity of the [^{57}FeNZr (4nm) / FeNZr (3nm)]$_{20}$ multilayer as a function of annealing time at 393 K. The peaks at $q_z \sim 0.5\,nm^{-1}$ are the "total reflection peaks" and those at $q_z \sim 0.9\,nm^{-1}$ the 1st order Bragg peaks. The latter diminish in time as the ^{57}Fe atoms diffuse through the multilayer structure. (right panel) Evolution of interdiffusion length L_d as a function of annealing time at different temperatures. The inset shows the Arrhenius behaviour of the diffusion coefficient with isothermal annealing temperature, yielding the activation energy and pre-exponent factor for self-diffusion of Fe (figure taken from [34]).*

be confirmed for NC $Fe_{90}Zr_7B_3$ disentangling the contributions from the grains and the grain boundaries. Most notably, the low-energy DOS exhibits a $g(E) = a \cdot E^2$ dependence over the whole range 1.5-15 meV for all measured samples (Figure 7 right panel (a)), and the coefficient a increases with decreasing d (Figure 7 right panel (b)). In summary, NC material seems to change properties not due to its grain size but due to the amount of grain and surface/boundary contributions.

Iron Self-diffusion in Nanocrystalline Alloys: With the newly developed method of nuclear resonance reflectivity using nuclear resonance scattering from isotopic multilayers the study of self-diffusion of Mössbauer isotopes over a wide range of diffusivity, $10^{-26}\,m^2s^{-1}$ to $10^{-18}\,m^2s^{-1}$, became feasible [33]. At the same time, this technique is sensitive enough to measure diffusion lengths as small as 0.1 nm, which is more than an order of magnitude larger than the sensitivity of conventional depth profiling techniques. Figure 8 (left panel) shows nuclear resonance reflectivity of isotopic multilayer of nanocrystalline FeNZr having structure [^{57}FeNZr (4nm)/FeNZr (3nm)]$_{20}$. The film is chemically homogeneous, however, scattering contrast between adjacent layers emanates due to the nuclear resonance scattering from the ^{57}Fe isotope. A Bragg peak around $q_z = 0.9\,nm^{-1}$ is observed due to isotopic periodicity. As the multilayer is annealed, iron atoms diffuse across the isotopic interfaces, and the intensity of the Bragg peak decays with annealing time, yielding the diffusion length through the relation, $\ln[I(t)/I(0)] = -4\pi^2 n^2 L_d^2/d^2$, where I(t) is the intensity of the Bragg peak after annealing time t, n is the order of Bragg reflection, L_d is the diffusion length and d is the bilayer periodicity. Figure 8 (right panel) gives L_d as a function of annealing time at three different temperatures T. It increases initially at a faster rate and after a certain annealing time the rate becomes constant. The initial faster increase in the diffusion length is associated with the structural relaxation in the system, similar

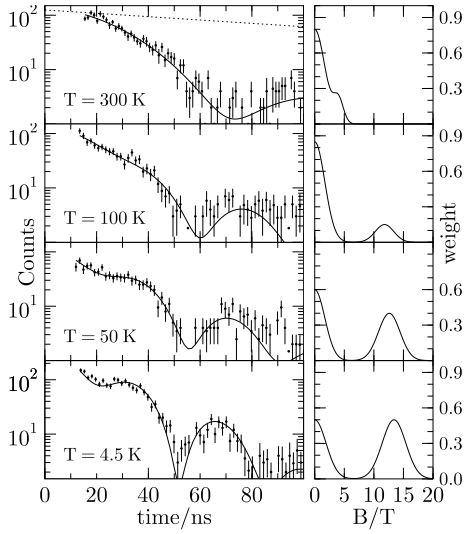


Fig. 9: *Time spectra of nuclear resonant grazing-incidence reflection from ultrasmall ^{57}Fe islands on W(110). The modulation of the intensity is attributed to a perpendicular magnetization of the Fe islands. The solid lines are the results of simulations. The right panel displays the probability density for the hyperfine field distribution that was obtained from the simulations (figure taken from [35]).*

to amorphous alloys. This relaxation is associated with the relaxation of the highly disordered grain boundary region, stress relaxation in the film and annealing out of the defects inside the nanocrystals. After sufficiently long annealing time, the diffusion length increases linearly with annealing time and the diffusivity was obtained by a linear fit using the relation: $L_d^2 = 2 \cdot D(T) \cdot t$. The diffusivity follows Arrhenius type behaviour with temperature, as shown in the inset of Figure 8 (right panel). The solid line in the figure is a fit to the experimental data obtained using the relation $D = D_0 \exp(-E/k_B T)$, where E is the activation energy and D_0 is the pre-exponential factor which contains the details of the diffusion mechanism. The activation energy for self-diffusion of Fe in the present system is significantly lower than that in the amorphous FeN of similar composition. This result is in general agreement with the earlier studies where the low activation energy for diffusion in nanocrystalline alloys has been attributed to a higher density of grain boundaries. Further, the observed values of D_0 and E suggest that the mechanism of diffusion in this system is similar to that in amorphous alloys.

4.2 Magnetism of nano-structured material

The understanding of magnetism of self organized nano-structures and clusters gained much attention because of fundamental aspects of nanoscale magnetic ordering in general and possible applications in high-density magnetic storage and magnetoelectronics.

Iron on W (110): As an example the perpendicular spin orientation in ultra small Fe islands on W (110) will be discussed [36]. Ultra small pseudomorphic Fe islands on an atomically clean W (110) crystal have been prepared by thermal evaporation of Fe enriched to 95 % in ^{57}Fe. The coverage of the Fe islands was 0.57 which is slightly below the percolation limit. Their average diameter was determined to be 2.0 nm. A coating of five monolayers of Ag prevents the sample from contamination.

Figure 9 (left panel) displays the time spectra taken in grazing incidence geometry between 300 K and 4.5 K. The right panel of Fig. 9 shows the weight of the magnetic component. It increases with decreasing temperature. This can be attributed to superparamagnetic relaxation of the magnetic moments. At high temperatures the magnetization of small particles is subject to fast thermal fluctuations so that the effective magnetic hyperfine field averages to zero. The transition from the fast relaxation regime to the magnetically ordered state occurs at about 50 K, however, it is a very broad transition certainly due to the size distribution of the islands. The modulation in the time spectra is characteristic for a perpendicular magnetization. This result is quite remarkable because Fe films on W (110) are known to be magnetized in plane for coverages of more than 0.6.

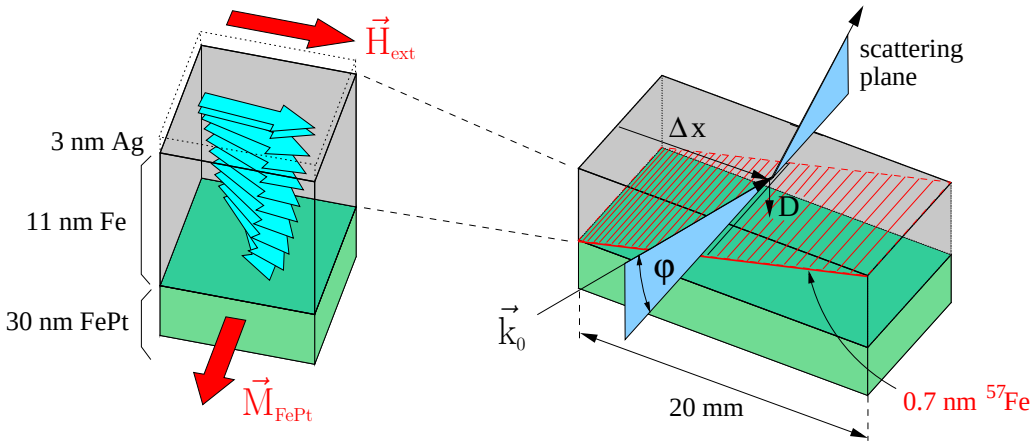


Fig. 10: *Magnetic structure in the iron layer of an Fe/FePt exchange spring as it forms in an external magnetic field $\vec{H}_{ext}$ of 160 mT. This image was obtained using an ultrathin probe layer embedded in the iron film, shown on the right (figure taken from [37]).*

Exchange-coupled thin films: Exchange-coupling between soft- and hard-magnetic phases plays an important role in the engineering of novel functional magnetic nanostructures. There exist some micromagnetical models describing this behaviour, however, due to methodological difficulties there exist nearly no direct measurements of the actual spin structure. Utilizing the technique of probe layers, i.e. inserting ^{57}Fe in various depths of the thin iron film (see Fig. 10 right panel), NRS directly probes the actual spin structure in various depths by selectively exciting the ^{57}Fe layer at various lateral positions [38]. The sample investigated here is a bilayer system consisting of 11 nm Fe on 30 nm $Fe_{55}Pt_{45}$ in the hard-magnetic tetragonal $L1_0$ phase. A wedge-shaped 0.7 nm thick ^{57}Fe film with a slope of 0.5 nm/mm has been produced. Different depths D in the sample can thus be probed by adjusting the displacement Δx of the sample transversely to the incident beam. The sample was magnetically saturated in an external field of 2.3 T so that the remanent magnetization $\vec{M}_{FePt}$ of the $Fe_{55}Pt_{45}$ layer was oriented along $\vec{k}_0$, as shown in Fig. 10. Time spectra at 4 K with selected in-plane external fields $\vec{H}_{ext}$ perpendicular to $\vec{k}_0$ were taken at various displacements Δx. Evaluation of the spectra reveals the depth dependence of the rotation of the magnetization in the iron film with atomic resolution (see Fig. 10 left panel).

5 Dynamics of disordered systems

NRS techniques are well suited for the investigation of the dynamics of disordered systems over a huge time scale. Whereas NIS covers the fast regime of phonons, NQES techniques complement the investigations to the slow time regime (10^5 - 10^{10} Hz). Furthermore, the technique of probe molecules allows for the investigation of glasses, which do not contain Mössbauer atoms and, when carefully chosen, allows for mode selectivity as collective and non-collective modes, respectively, and translational and rotational relaxations, respectively.
The following two examples cover both aspects, fast dynamics of glasses with emphasis to the so-called "boson peak" and the nature of α- and slow β-relaxation in the glass former dibutyl phthalate (DBP).

5.1 Universal dynamics of glass

Glass dynamics has been studied since decades and the boson peak more than 30 years. However, the nature of it is still debated. NIS allowed for the first time to study selectively collective modes utilizing the technique of "probe molecules". The probes, which are molecules with a resonant (Mössbauer) nucleus in the centre of mass "swims" in a glass matrix without coupling to the host molecules. It follows the collective motions of the glass without being sensitive to the local modes of the host molecule. Furthermore, one selects pure translational motions of the probes: rotation is disregarded because the spectator nucleus is in the centre of mass; the few intra-molecular modes are separated in energy. The selected pure translational motions of the probe give the "density of states of collective motions" (CDOS) of the glass matrix.

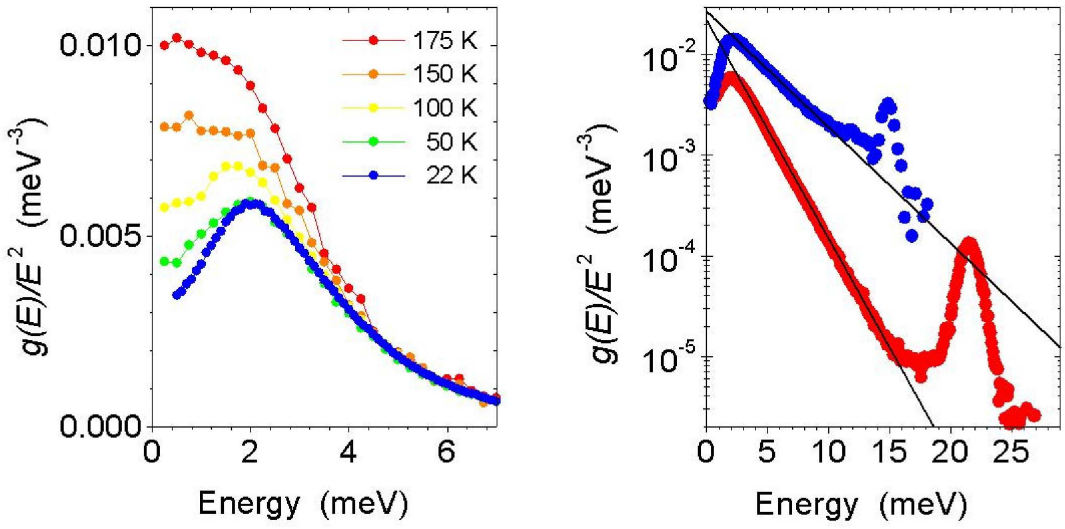


Fig. 11: *left panel: Temperature dependence of the reduced DOS of collective motions of ferrocene in dibutyl phthalate. right panel: Reduced DOS of collective motions of ferrocene in DBP from NIS measurements (lower curve, red) and of pure DBP from neutron mesurements (upper curve, blue). The neutron data are scaled to match the NIS data at lowest energy. Solid lines show the fit according to the equation (see text). (figure taken from [39])*

For all studied glasses, the reduced CDOS g(E)/E^2 clearly exhibits an excess of low-energy modes, the boson peak, which is shown exemplarily for ferrocene in DBP in Fig. 11 (left panel).

The positions of the peak are consistent with the boson peak energies in the total DOS from neutron scattering data. The temperature evolution of the boson peak shows the same features as observed with other methods. This clear manifestation of the boson peak in the CDOS proves that the boson peak in the total DOS is largely composed of collective modes. Beyond the boson peak, the reduced density of states of collective motions reveals a temperature-independent exponential behaviour: $g(E)/E^2 \propto \exp(-E/E_0)$, which was not reported earlier. On a logarithmic scale, $g(E)/E^2$ follows a straight line over three decades of the reduced CDOS and deviates only when obscured by the eigen mode of ferrocene at 22 meV (Figure 11 right panel). We also found this exponential behaviour in the total DOS available from neutron data [40]. Further investigations towards the understanding of the boson peak have been carried out on hyperquenched [41] and on densified glasses [42]. All these results suggest that for a glass with a given microscopic structure the transformation of atomic vibrations is sufficiently described by the corresponding transformation of the elastic medium. Finally, a comparison of the boson peak with the transverse acoustic van Hove singularity in the corresponding crystal suggests the identical nature of the two features [43].

5.2 α- and β-relaxation

Approaching the slow time regime relaxation times and transport properties of moderately supercooled liquids exhibit a universal dependence on temperature. At the glass transition temperature, T_g, the relaxation of organic glass formers usually splits into two processes, the α-relaxation, which corresponds to cooperative molecular dynamics, and the slow β-relaxation, which microscopic origin is still disputed. NRS techniques may shine some more light in the underlaying mechanism. The above studied DBP with the probe molecule ferrocene (T_g = 178 K) was investigated by NFS (sensitive to translational and rotational dynamics) and SRPAC (sensitive only to rotational dynamics) [44].

NFS spectroscopy is applicable as long as the Lamb-Mössbauer factor f_{LM} is greater than zero which holds in this case up to 210 K. SRPAC on the other hand does not depend on f_{LM} and can be applied in the entire temperature regime. At the lower temperatures, the SRPAC intensity follows an exponential decay modulated by quantum beats. In the regime of slow relaxation, the beats are damped at a rate proportional to rotational relaxation. Similar quantum beats modulate the decay of the NFS intensity, where the damping depends on the sum of rotational and translational relaxation. At higher temperatures, in the regime of fast relaxation, only SRPAC spectra can be measured. From the measured time spectra relaxation rates have been determined (see Figure 12).

Below 190 K the data sets from both techniques coincide, which means that translational dynamics is absent in the experimental time window. Above 190 K the NFS data begin to deviate from the SRPAC data because translational dynamics is activated. From these results the pure translational relaxation rates have been derived. Comparing the NRS relaxation rates with data of pure DBP as obtained from dielectric spectroscopy (DS), we may now interpret those two branches in DS at low temperature. The branch of slow β-relaxation follows the NRS data of rotational dynamics, whereas the branch of α-relaxation decreases in parallel with the NRS data of translational dynamics. This correlation suggest a decoupling and, especially, that the slow β-relaxation is connected with rotational dynamics.

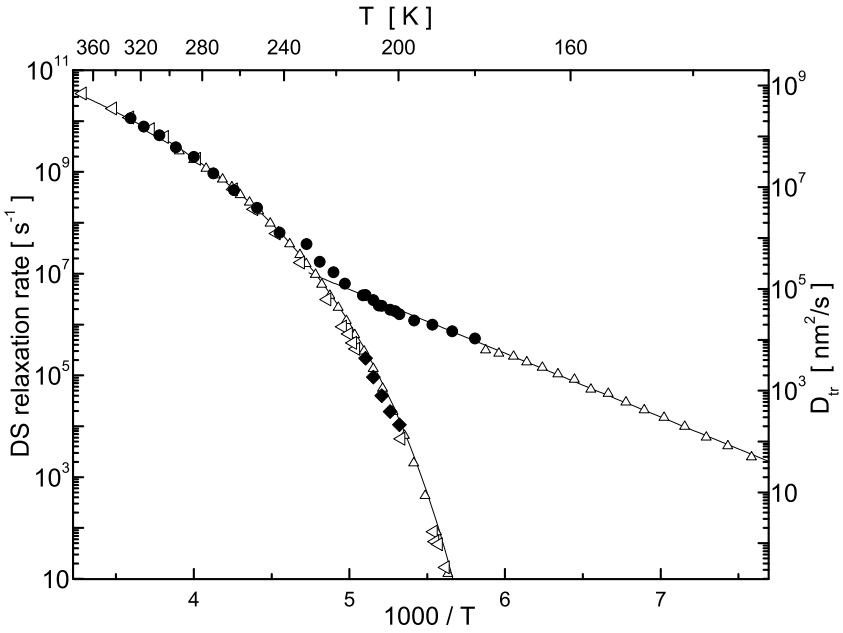


Fig. 12: *Comparison of the rotational relaxation rate of the ferrocene molecule in DBP (glass temperature T_G = 178 K) derived from SRPAC (•) and the translational relaxation rate derived from NFS (⧫) to the DS data [45] for pure DBP (▷). (figure taken from [46]).*

6 Further applications and summary

In the previous chapters selected topics have been presented. Others could not be mentioned as investigations of biological relevant samples which became a growing field with emphasis to their dynamic and electronic properties (see e.g. [47, 48, 49]), thermoelectric materials [50, 51], or geoscience [52, 53]. The outstanding energy resolution and the timing properties of NRS led to applications beyond spectroscopy. Contributions to the definition of the length scale [54], correlation behaviour [55], interferometry [56, 57], the utilization of waveguides [58], and the contribution to quantum electrodynamics [59] show the hole breath of the method. All those applications became possible with the advent of powerful synchrotron radiation sources around the world as APS (Argonne, USA), SPring-8 (Harima, Japan), ESRF (Grenoble, France), and, just coming, PETRA III (Hamburg, Germany). Nowadays nuclear resonance scattering techniques matured to a spectroscopy with unique applications in various fields. The probe molecule/layer technique, the application of sophisticated sample environments, and the possibility to apply simultaneously various NRS techniques and other techniques makes this spectroscopy an unique tool. The previous typical examples demonstrate the versatility and huge potential of the technique.

References

[1] E. Gerdau, R. Rüffer, H. Winkler, W. Tolksdorf, C. P. Klages, and J. P. Hannon, Phys. Rev. Lett. **54**, 835 (1985).

[2] E. Gerdau and H. de Waard, eds., *Nuclear Resonant Scattering of Synchrotron Radiation*, vol. 123-125 of *Hyperfine Interactions* (Baltzer Science Publishers, 1999/2000).

[3] R. Röhlsberger, *Nuclear Condensed Matter Physics with Synchrotron Radiation - Basic Principles, Methodology and Applications*, vol. 208 of *Springer Tracts in Modern Physics* (Springer Publishers, 2004).

[4] J. B. Hastings, D. P. Siddons, U. van Bürck, R. Hollatz, and U. Bergmann, Phys. Rev. Lett. **66**, 770 (1991).

[5] U. van Bürck, D. P. Siddons, J. B. Hastings, U. Bergmann, and R. Hollatz, Phys. Rev. B **46**, 6207 (1992).

[6] H. Grünsteudel, *Nuclear Resonant Scattering of Synchrotron Radiation on Iron Containing Biomimetic Compounds*, Ph.D. thesis, Medizinische Universiät zu Lübeck, Lübeck (1998).

[7] O. Leupold, K. Rupprecht, and G. Wortmann, Structural Chemistry **14**, 97 (2003).

[8] S. Dattagupta, Physical Review B **14**, 1329 (1976).

[9] M. Seto, Y. Yoda, S. Kikuta, X. W. Zhang, and M. Ando, Phys. Rev. Lett. **74**, 3828 (1995).

[10] W. Sturhahn, T. S. Toellner, E. E. Alp, X. Zhang, M. Ando, Y. Yoda, S. Kikuta, M. Seto, C. W. Kimball, and B. Dabrowski, Phys. Rev. Lett. **74**, 3832 (1995).

[11] A. I. Chumakov, R. Rüffer, H. Grünsteudel, H. F. Grünsteudel, G. Grübel, J. Metge, O. Leupold, and H. Goodwin, Europhysics Letters **30**, 427 (1995).

[12] A. I. Chumakov and W. Sturhahn, Hyperfine Interactions **123/124**, 781 (1999).

[13] K. Singwi and A. Sjölander, Physical Review **120**, 1093 (1960).

[14] V. G. Kohn, A. I. Chumakov, and R. Rüffer, Physical Review B **58**, 8437 (1998).

[15] W. Sturhahn and V. G. Kohn, Hyperfine Interactions **123/124**, 367 (1999).

[16] H. J. Lipkin, Hyperfine Interactions **123/124**, 349 (1999).

[17] J. Zhao, W. Sturhahn, J. Lin, G. Shen, E. E. Alp, and H. Mao, High Pressure Research **24**, 447 (2004).

[18] J. Lin, W. Sturhahn, J. Zhao, G. Shen, H. Mao, and R. Hemley, Geophysical Research Letters **31**, L14611 (2004).

[19] H. F. Grünsteudel, *Der α-ϵ-Übergang in Eisen als Beispiel für nukleare Vorwärtsstreuung von Synchrotronstrahlung an Proben unter hohem Druck*, Ph.D. thesis, Universiät Paderborn, Paderborn (1997).

[20] R. Lübbers, H. F. Grünsteudel, A. I. Chumakov, and G. Wortmann, Science **287**, 1250 (2000).

[21] H. K. Mao, J. Xu, V. V. Struzhkin, J. Shu, R. J. Hemley, W. Sturhahn, M. Y. Hu, E. E. Alp, L. Vocadlo, D. Alfé, G. D. Price, M. J. Gillan, *et al.*, Science **292**, 914 (2001).

[22] R. Lübbers, H. Giefers, K. Rupprecht, G. Wortmann, and A. I. Chumakov, ESRF Highlights 2000 pp. 48–49 (2001).

[23] H. Giefers, R. Lübbers, K. Rupprecht, G. Wortmann, D. Alfe, and A. Chumakov, High Pressure Research **22**, 501 (2002).

[24] D. Antonangeli, F. Occelli, H. Requardt, J. Badro, G. Fiquet, and M. Krisch, Earth and Planetary Science Letters **225**, 243 (2004).

[25] J. Brown and R. McQueen, Journal of Geophysical Research **91**, 7485 (1986).

[26] A. Barla, J. P. Sanchez, Y. Haga, G. Lapertot, B. P. Doyle, O. Leupold, R. Rüffer, M. M. Abd-Elmeguid, R. Lengsdorf, and J. Flouquet, ESRF Highlights 2003 pp. 6–7 (2004).

[27] A. Barla, J. Derr, J. Sanchez, B. Salce, G. Lapertot, B. Doyle, R. Rüffer, R. Lengsdorf, M. Abd-Elmeguid, and J. Flouquet, ESRF Highlights **2005**, 9 (2005).

[28] A. Barla, J. P. Sanchez, Y. Haga, G. Lapertot, B. P. Doyle, O. Leupold, R. Rüffer, M. M. Abd-Elmeguid, R. Lengsdorf, and J. Flouquet, Phys. Rev. Lett. **92**, 066401 (2004).

[29] E. Annese, A. Barla, C. Dallera, G. Lapertot, J.-P. Sanchez, and G. Vankó, Phys. Rev. B **73**, 140409 (2006).

[30] L. Pasquini, A. Barla, A. I. Chumakov, O. Leupold, R. Rüffer, A. Deriu, and E. Bonetti, ESRF Highlights 2002 pp. 57–58 (2003).

[31] L. Pasquini, A. Barla, A. I. Chumakov, O. Leupold, R. Rüffer, A. Deriu, and E. Bonetti, Phys. Rev. B **66**, 073410 (2002).

[32] S. Stankov, Y. Z. Yue, M. Miglierini, B. Sepiol, I. Sergueev, A. I. Chumakov, L. Hu, P. Svec, and R. Rüffer, Physical Review Letters **100**, 235503 (2008).

[33] A. Gupta, M. Gupta, S. Chakravarty, R. Rüffer, H.-C. Wille, and O. Leupold, Phys. Rev. B **72**, 014207 (2005).

[34] A. Gupta, M. Gupta, S. Chakravarty, R. Rüffer, H.-C. Wille, and O. Leupold, ESRF Highlights 2005 pp. 8–9 (2006).

[35] R. Röhlsberger, J. Bansmann, V. Senz, K.-L. Jonas, A. Bettac, O. Leupold, R. Rüffer, E. Burkel, and K.-H. Meiwes-Broer, ESRF Highlights 2001 p. 69 (2002).

[36] R. Röhlsberger, J. Bansmann, V. Senz, K. L. Jonas, A. Bettac, O. Leupold, R. Rüffer, E. Burkel, and K. H. Meiwes-Broer, Phys. Rev. Lett. **86**, 5597 (2001).

[37] T. Klein, R. Röhlsberger, K. Schlage, H. Thomas, O. Leupold, and E. Burkel, ESRF Highlights 2004 pp. 11–12 (2005).

[38] R. Röhlsberger, H. Thomas, K. Schlage, E. Burkel, O. Leupold, and R. Rüffer, Phys. Rev. Lett. **89**, 237201 (2002).

[39] A. I. Chumakov, I. Sergueev, U. van Bürck, W. Schirmacher, T. Asthalter, R. Rüffer, O. Leupold, and W. Petry, ESRF Highlights 2004 pp. 7–8 (2005).

[40] A. I. Chumakov, I. Sergueev, U. van Bürck, W. Schirmacher, T. Asthalter, R. Rüffer, O. Leupold, and W. Petry, Phys. Rev. Lett. **92**, 245508 (2004).

[41] A. Monaco, A. I. Chumakov, Y.-Z. Yue, G. Monaco, L. Comez, D. Fioretto, W. A. Crichton, and R. Rüffer, Phys. Rev. Lett. **96**, 205502 (2006).

[42] A. Monaco, A. I. Chumakov, G. Monaco, W. A. Crichton, A. Meyer, L. Comez, D. Fioretto, J. Korecki, and R. Rüffer, Phys. Rev. Lett. **97**, 135501 (2006).

[43] A. I. Chumakov, G. Monaco, A. Monaco, W. Crichton, A. Bosak, R. Rüffer, A. Meyer, F. Kargl, L. Comez, D. Fioretto, H. Giefers, S. Roitsch, *et al.*, Phys. Rev. Lett. **106**, 225501 (2011).

[44] I. Sergueev, U. van Bürck, A. I. Chumakov, T. Asthalter, G. V. Smirnov, H. Franz, R. Rüffer, and W. Petry, Phys. Rev. B **73**, 024203 (2006).

[45] M. Sekula, S. Pawlus, S. Hensel-Bielowka, J. Ziolo, M. Paluch, and C. M. Roland, J. Phys. Chem. B **108**, 4997 (2004).

[46] I. Sergueev, U. van Bürck, A. I. Chumakov, T. Asthalter, G. V. Smirnov, H. Franz, R. Rüffer, and W. Petry, ESRF Highlights 2003 pp. 12–13 (2004).

[47] J. T. Sage, S. M. Durbin, W. Sturhahn, D. C. Wharton, P. M. Champion, P. Hession, J. Sutter, and E. E. Alp, Phys. Rev. Lett. **86**, 4966 (2001).

[48] K. Achterhold, C. Keppler, A. Ostermann, U. van Bürck, W. Sturhahn, E. E. Alp, and F. G. Parak, Phys. Rev. E **65**, 051916 (2002).

[49] H. Winkler, A. Chumakov, and A. Trautwein, in *Spin Crossover in Transition Metal Compounds III*, edited by P. Gütlich and H. Goodwin (Springer, 2004), vol. 235 of *Topics in Current Chemistry*, pp. 137–152.

[50] G. J. Long, R. P. Hermann, F. Grandjean, E. E. Alp, W. Sturhahn, C. E. Johnson, D. E. Brown, O. Leupold, and R. Rüffer, Physical Review B **71**, 140302 (2005).

[51] S. Tsutsui, J. Umemura, H. Kobayashi, Y. Yoda, H. Onodera, H. Sugawara, D. Kikuchi, H. Sato, C. Sekine, and I. Shirotani, Physica B-Condensed Matter **383**, 142 (2006).

[52] C. McCammon, I. Kantor, O. Narygina, J. Rouquette, U. Ponkratz, I. Sergueev, M. Mezouar, V. Prakapenka, and L. Dubrovinsky, Nature Geoscience **1**, 684 (2008).

[53] L. Dubrovinsky, K. Glazyrin, C. McCammon, O. Narygina, E. Greenberg, S. Übelhack, A. I. Chumakov, S. Pascarelli, V. Prakapenka, J. Bock, and N. Dubrovinskaia, Journal of Synchrotron Radiation **16**, 737 (2009).

[54] Y. V. Shvyd'ko, M. Lerche, J. Jäschke, M. Lucht, E. Gerdau, M. Gerken, H. D. Rüter, H.-C. Wille, P. Becker, E. E. Alp, W. Sturhahn, J. Sutter, *et al.*, Phys. Rev. Lett. **85**, 495 (2000).

[55] E. Gluskin, E. Alp, I. McNulty, W. Sturhahn, and J. Sutter, Journal of Synchrotron Radiation **6**, 1065 (1999).

[56] Y. V. Shvyd'ko, M. Lerche, H.-C. Wille, E. Gerdau, M. Lucht, H. D. Rüter, E. E. Alp, and R. Khachatryan, Phys. Rev. Lett. **90**, 013904 (2003).

[57] M. Yabashi, K. Tamasaku, and T. Ishikawa, Phys. Rev. Lett. **88**, 244801 (2002).

[58] R. Röhlsberger, K. Schlage, T. Klein, and O. Leupold, Phys. Rev. Lett. **95**, 097601 (2005).

[59] R. Röhlsberger, K. Schlage, B. Sahoo, S. Couet, and R. Rüffer, Science **328**, 1248 (2010).

D 10 Coherent Imaging with X-ray Free-Electron Lasers[1]

H. N. Chapman
Center for Free-Electron Laser Science
DESY, Hamburg

Contents

[1]Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

X-ray free-electron lasers (FELs) such as the Linac Coherent Light Source (LCLS) [1] produce laser-like pulses of X-rays of 10 fs to 300 fs duration and up to about 10^{13} photons per pulse. Just as with the introduction of the synchrotron, X-ray FELs promise to have a great impact on many scientific disciplines by opening up the study of materials at the length scale of interatomic distances and at the corresponding time scales of atomic motion. In condensed matter science, for example, materials properties can be drastically altered by inducing transient structures using ultrafast light pulses [2]. Similarly, matter at extreme limits of temperature and pressure, similar to conditions in the cores of stars and planets, can be created transiently in the laboratory using intense optical pulses. The extremely intense and brief pulses delivered by X-ray FELs opens up the direct imaging of processes in these systems, which previously could only be studied by spectroscopic means. In the biological sciences, time-resolved X-ray crystallography has elucidated the atomic motions that occur on stimulation of light-sensitive proteins [3]. This field will experience a dramatic increase of capability with FEL sources. The extreme irradiance[2] of the pulses will let us shrink crystal sizes all the way down to single molecules, giving three-dimensional movies of conformational dynamics and chemical reactions, and allowing the imaging of macromolecules that cannot be easily crystallized.

X-ray crystallography is indeed the inspiration for the methods of imaging at X-ray FELs [4]. Crystallography can obtain atomic-resolution images of protein macromolecules without the use of a lens, based on the interpretation of the coherent scattering pattern detected in the far field. Since this method is lensless, there are no technological limitations to the achievable resolution caused by the perfection of objective lenses that can be made. Instead, the image is synthesized from the measured diffraction intensities. Each intensity sample is related to the strength of each spatial frequency component of the synthesized image. The shift of each spatial frequency component in real space is given by the corresponding phase of the diffracted wavefield. Crucially, this information cannot be obtained by measurement. The correct image, formed by the sum of the correctly positioned periodic components, cannot be synthesized without this information. This so-called phase problem has been largely solved in protein crystallography by constraining the phases through additional information about the structure such its atomicity (if scattering to high enough resolution was obtained), protein sequence, or known similar structure. Anomalous diffraction methods can be used to determine the locations of heavy atoms, which in turn can act as holographic references to obtain the remaining image. This general imaging strategy works even better for non-crystalline materials, since the information content of a continuous diffraction pattern of a non-periodic object is large enough to completely constrain the phases, allowing *ab initio* reconstruction of both two- and three-dimensional images [5, 6]. The phase retrieval algorithms in this case were first developed in the context of electron microscopy [7] and optical remote sensing [8].

The generalization of crystallography tol non-periodic objects requires sufficient spatial and temporal coherence of the illumination wavefield so that the scattered waves from extreme points in the object interfere to create the fringe pattern on the detector that encodes their separation. The lack of crystallinity does come at a severe cost of diffraction signal, since the integrated Bragg intensity is proportional to the number of units in the crystal [9] and the number of repeats in even small protein crystals can number in the billions. Unfortunately the lower

[2]Pulse irradiance is given in units of photons per unit area and time. It is often referred to as "intensity", although in crystallography that word commonly means "diffracted counts" which is the meaning employed in this chapter.

signal cannot simply be compensated by longer exposure, since biological materials can only withstand a limited dose before they are destroyed by the very beam that is used to obtain the image [10]. But here the ultrafast pulses of X-ray FELs come to the rescue, by diffracting from the sample before the effects of radiation damage have set in [11]. There is strong evidence that this concept of "diffraction before destruction" holds to atomic resolution, at doses over 100 times higher than can be tolerated by slow exposures [12]. The development of this imaging technique is following two tracks: application to unique structures, such as cells or soot particles, that can only be imaged in a single shot; and to reproducible objects, such as viruses, macromolecular complexes, and protein nanocrystals, and where signals can be accumulated over many copies in a rapid series of diffraction measurements. These are extremely active fields of research, and progress is fast despite the fact that only a single hard X-ray FEL facility currently exists.

In this lecture, a brief description of coherent scattering and iterative phasing techniques are described, showing how 2D and 3D images of finite-size non-periodic objects can be recovered from diffraction data. Examples of the application of this technique are presented from experiments carried out at the FLASH soft X-ray FEL in Hamburg and the LCLS, and the research of single-particle diffractive imaging and "diffraction before destruction" is summarized, including computational and experimental investigations of the rate of perturbation of a structure under intense X-ray illumination.

2 Coherent Scattering from Isolated Objects

The interaction of x-rays with matter can be described by the inhomogeneous Helmholtz equation,

$$(\nabla^2 + k^2)\psi(\mathbf{x}) = \Phi(\mathbf{x})\psi(\mathbf{x}) \tag{1}$$

for an incident plane wave of wavelength λ or wavenumber $k = 2\pi/\lambda$. The scattering potential of a static and non-magnetic object may be expressed as

$$\Phi(\mathbf{x}) = k^2(1 - n^2(\mathbf{x})) = 4\pi r_e \rho(\mathbf{x}) \tag{2}$$

where n is the refractive index, r_e the classical radius of the electron, and $\rho(\mathbf{x})$ the electron density of the object. In the Born approximation the solution to (1) is of the form

$$\psi(\mathbf{x}) = e^{i\mathbf{k}_{\rm in}\cdot\mathbf{x}} + \frac{e^{ikr}}{r} f(\mathbf{q}) \tag{3}$$

where

$$f(\mathbf{q}) = -r_e \int \rho(\mathbf{x}) \exp(i\mathbf{q}\cdot\mathbf{x})\, d\mathbf{x}. \tag{4}$$

and $\mathbf{q} = \mathbf{k}_{\rm out} - \mathbf{k}_{\rm in}$ is the photon momentum transfer. The physical picture that explains the form of the scattering factor $f(\mathbf{q})$ is shown in Fig. 1 (a). A ray scattered in a direction $\mathbf{k}_{\rm out}$ from a point $\mathbf{x}_1$ will acquire a path difference of $\ell_1 = (\mathbf{x}_1 \cdot \hat{\mathbf{k}}_{\rm out} - \mathbf{x}_1 \cdot \hat{\mathbf{k}}_{\rm in})$ relative to a ray scattering from the origin $\mathbf{O}$, where $\hat{\mathbf{k}}$ are unit vectors. This is the difference of the lengths of the thick lines in Fig. 1 (a). The accumulated phase will therefore be $\phi_1 = (2\pi/\lambda)\ell_1 = \mathbf{x}_1 \cdot \mathbf{q}$. The point scatterer itself may cause a modification to the wave by the complex constant f_1, giving a scattering $f_1 \exp(i\phi_1) = f_1 \exp(i\mathbf{x}_1 \cdot \mathbf{q})$. Equation (4) is simply the integration over all scatterers in the object. The object's electron density is assumed to be continuous, although it can certainly be atomistic and may possess translational periodicity (i.e. be a crystal).

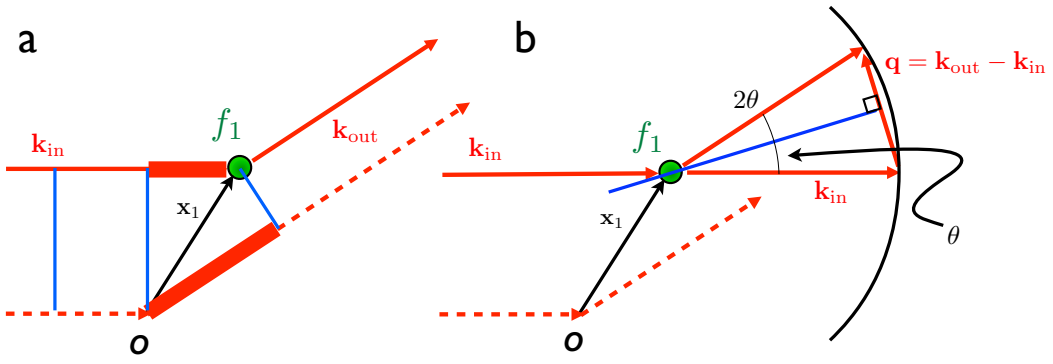


Fig. 1: *The far-field scattering geometry and the Ewald-sphere construction.*

Equation (4) states that the scattering amplitude f is given by the Fourier transform of the electron density. That is, the strength of diffraction in the direction $\mathbf{k}_{\rm out}$ only depends on the Fourier component $\tilde{\rho}(\mathbf{q})$, where we define the Fourier transform as

$$\tilde{g}(\mathbf{q}) \equiv \mathcal{F}\{g(\mathbf{x})\} \equiv \int g(\mathbf{x}) \exp(i\mathbf{x}\cdot\mathbf{q})\, d\mathbf{x}. \tag{5}$$

This component is a particular spatial frequency in the object, which may be thought of as a volume grating of a particular wavenumber $|\mathbf{q}|$ and direction $\hat{\mathbf{q}}$. From Fig. 1 (b), it is seen that the magnitude of $\mathbf{q}$ is given simply by

$$|\mathbf{q}| = 2|\mathbf{k}| \sin\theta = \frac{4\pi}{\lambda} \sin\theta \tag{6}$$

for a scattering angle 2θ, and that due to the conservation of $\mathbf{k}$ (that is, elastic scattering) the vector $\mathbf{q}$ lies on the surface of a sphere (called the Ewald sphere). We see from the diagram in Fig. 1 (b) that the scattered ray appears to reflect at an angle θ from a plane normal to $\mathbf{q}$. That is, the ray reflects from the volume grating which is tilted at the angle θ relative to the incoming wave-vector. The ray only reflects if the period of the volume grating, $d = 2\pi/|\mathbf{q}|$ satisfies Eqn. (6), which is to say $d = \lambda/(2\sin\theta)$ which is well recognized as Bragg's law. We stress that although Bragg's law and the Ewald sphere construction are well known concepts in crystallography, there is no requirement of periodicity of the object in the derivation or application of these concepts.

A typical coherent diffraction experiment illuminates a sample of electron density $\rho(\mathbf{x})$ (or refractive index $n(\mathbf{x})$) with a quasi-monochromatic plane wave and measures the diffraction pattern in the far field with a planar detector (such as a bare CCD). This is a measure of the intensity of the wavefield, given by

$$I(\mathbf{q}) = I_0\, \Omega_p\, P\, r_e^2\, |\tilde{\rho}(\mathbf{q})|^2\,, \tag{7}$$

for pixels of solid angle Ω_p, and where P is the polarization factor. The mapping from pixel coordinate to $\mathbf{q}$ is easily obtained by the geometry shown in Fig. 1 where $\mathbf{k}_{\rm out}$ points in the direction of the pixel. Note that in a single exposure in this geometry records only information about the object for spatial frequencies that lie on the Ewald sphere. Other frequencies are missing in the measurement and the only way to record the full 3D information is to record

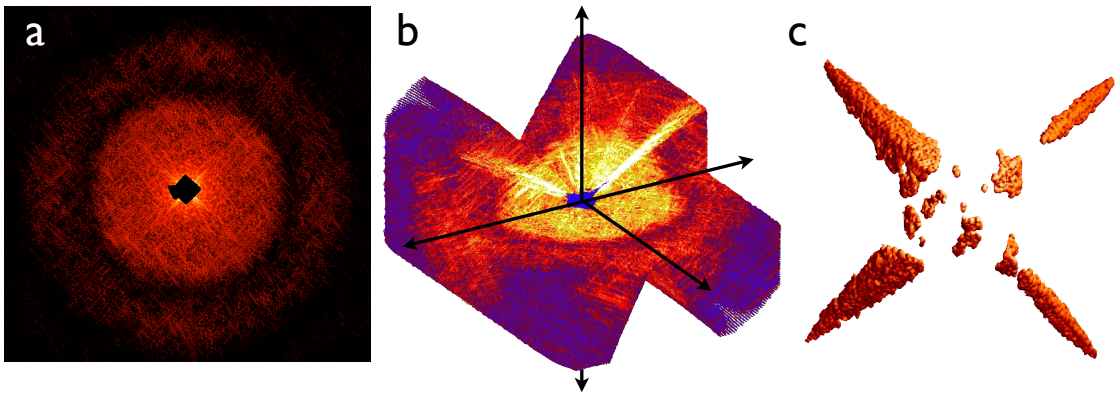


Fig. 2: *Diffraction data collected from a 3D test object, showing (a) a diffraction pattern recorded at a single orientation, (b) 3D diffraction intensities collected at orientations from* $-70°$ *to* $+70°$*, and (c) the reconstructed volume image [6].*

additional diffraction patterns for different orientations of the object relative to the incident beam. Rotating the sample rotates its Fourier spectrum such that it sweeps across the Ewald sphere. From the point of view of the object, the measurements at different orientations can be thought of as orienting the Ewald sphere at different rotations (about the point $\mathbf{q} = 0$) so that it samples the 3D intensities $|\tilde{\rho}(\mathbf{q})|^2$. An example of 3D diffraction data collected at a synchrotron beamline is shown in Fig. 2.

Since the sample is completely destroyed by the interaction with a single high-irradiance FEL pulse, one object can only give rise to a single measurement (or possibly multiple simultaneous measurements). Full 3D information therefore requires combining many diffraction patterns, each from a different orientation of an identical copy of the object. Combining data from many patterns is also needed to increase the overall signal. A calculated single molecule diffraction pattern is shown in Fig. 3 for an incident pulse fluence of 10^{12} photons focused to a 0.3-μm diameter spot at 8 keV photon energy ($10^6\,\mathrm{J/cm^2}$), corresponding to an irradiance of $3 \times 10^{19}\,\mathrm{W/cm^2}$ with a pulse duration of 30 fs. The pattern is noisy because of low photon counts (most detector pixel values are zero or one photon count). Therefore, even with the extremely high irradiances from an X-ray FEL, some averaging of the signal from many particles is required to increase the signal to noise ratio even at a single orientation. The photon count per pixel of solid-angle Ω_p per shot, averaged over shells of q, for biological material can be estimated by

$$\langle I(q)\rangle_q = I_0\,\Omega_p\,P\,r_e^2\,\left\langle|\tilde{\rho}(q)|^2\right\rangle_q = I_0\,\Omega_p\,P\,r_e^2\,N_{\mathrm{atom}}\,|f|^2, \tag{8}$$

where N_{atom} are the number of atoms in the molecule, and f is an average atomic scattering factor (e.g. close to that of carbon). Full diffraction information requires sampling at the Shannon rate Δq_S as described in Sec. 3.1 on p. 10.9. In this case we have $\Omega_p \approx (\lambda/4\pi)^2\Delta q_S^2 = (\lambda/4w)^2$ for a particle of width w, or

$$\langle I(q)\rangle_q = I_0\,P\,r_e^2\,\frac{N_{\mathrm{atom}}\lambda^2}{16w^2}\,|f|^2 \tag{9}$$

in the case where there is no sample motion during the pulse. Equations (8) and (9) assume that the positions of atoms are completely uncorrelated, which is approximately true at resolutions approaching the atomic scale. They give an estimate of average counts per pixel per particle for

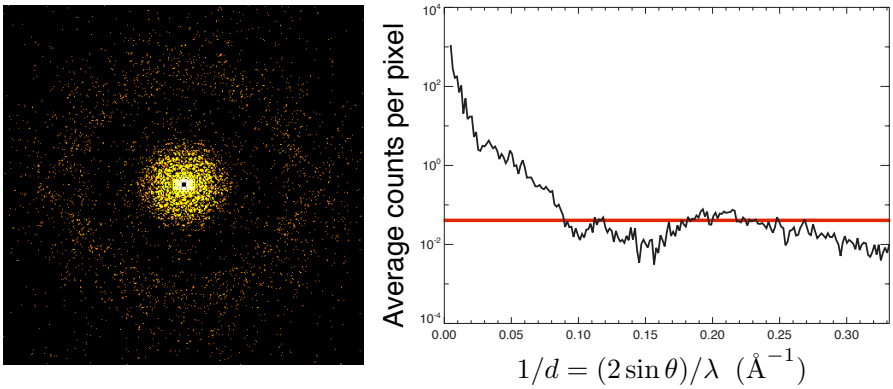


Fig. 3: *Calculated diffraction pattern of a single cow pea mosaic virus (CPMV) particle for* 10^{12} *incident photons focused to a* 0.3*-*μm *diameter spot, at 8 keV photon energy.*

a single shot. For the CPMV particle, of width $w = 31\,\mathrm{nm}$, we obtain $\langle I(q)\rangle = 0.04$ photons per Shannon pixel, indicated by the red line in the plot of Fig. 3. We see from Eqn. (9) that there is a large incentive to perform measurements at as long a wavelength as possible that can support the desired resolution, and that the scattering signal is not strongly dependent on the size of the object, since $N_{\mathrm{atom}} \propto w^3$. With this model we find that biological particles of around 10 nm diameter require a pulse fluence of 10^{13} photons/$(0.1\mu\mathrm{m})^2$ for an average of 0.1 photon per Shannon pixel at a resolution of 3 Å and a photon energy of 8 keV. For a pulse duration of 30 fs this requires the exceptionally high irradiance of $4 \times 10^{21}\,\mathrm{W/cm^2}$, the effect of which is examined in Sec. 4. At 3 keV photon energy, the required irradiance is reduced to $2 \times 10^{20}\,\mathrm{W/cm^2}$, although the increased photoabsorption cross section at this photon energy leads to faster destruction of the object. The signal level of 0.1 photon per Shannon pixel is higher than required by the averaging and assembly procedures discussed in Sec. 5.

3 Image Reconstruction

3.1 The Phase Problem

The reconstruction of a 2D or 3D image of the electron density $\rho(\mathbf{x})$ requires the inversion of Eqn. (7). While the modulus of the Fourier amplitudes can be obtained from $\sqrt{I}(\mathbf{q})$, the phases are missing. That is, we know the strength $|\tilde{\rho}|$ of each volume grating (of a specified period and direction given by $\mathbf{q}$) in the image, but not how these gratings are shifted with respect to each other. If the phases are known then the synthesized image is a coherent sum of these properly positioned frequency components, which is simply the inverse Fourier transform of $\tilde{\rho}(\mathbf{q})$:

$$i(\mathbf{x}) = \mathcal{F}^{-1}\{\tilde{\rho}(\mathbf{q})\} \equiv \int \tilde{\rho}(\mathbf{q}) \exp(-i\mathbf{q}\cdot\mathbf{x})\, d\mathbf{q}. \tag{10}$$

In the case of reconstructing an image from a single diffraction pattern, this image synthesis is the same as would be carried out by a perfect lens. The lens of course avoids the need to retrieve phases, since it acts upon the far-field wavefield incident on its pupil, and in a microscope the image intensity is directly measured, not the far-field pattern. The resolution of the lens-based image is equivalent to that retrieved from the diffraction pattern if the acceptance of the lens is

the same as the angular extent of the recorded diffraction pattern. For example, a diffraction pattern that extends to a scattering angle of 2θ will resolve spatial periods in the object as small as $d = \lambda/(2\sin\theta)$. Similarly the lens-based image (of a coherently illuminated object) has a resolution commonly expressed as $d = \lambda/\mathrm{NA} = \lambda/\sin(2\theta)$, where NA is the numerical aperture of the lens.

Note that just having a detector of large angular extent does not necessarily guarantee a high-resolution image. The high-angle diffraction data of course must be detected above noise. The same is indeed true for a lens-based microscope, where the resolution of the image might not fulfill that expected by the numerical aperture of the lens. An estimation of the resolution in an X-ray or electron microscope is obtained by examining the Fourier spectrum of the recorded image. For our "lensless" imaging method we can look directly at the recorded diffraction pattern, although the resolution of the image will depend upon the reliability of the phases obtained. Methods of estimating this from the reproducibility of retrieved phases have been developed [13], but this does not necessarily mean the phases are the correct ones. Perhaps the most satisfactory way of estimating resolution is based on comparing images (or diffraction phases) retrieved in separate measurements, as is common in crystallography and single-particle electron microscopy.

The coherent X-ray diffraction data in Fig. 2 contains $512 \times 512 \times 512$ pixels, or measurements of Fourier amplitudes, $\sqrt{I}(\mathbf{q})$. Inversion to an image in this case requires the recovery of over 10^8 phases. The retrieval of these phases depends upon the observation that the number of independent measurements may exceed the number of degrees of freedom describing the image. A much simpler example is the diffraction pattern of a pair of point scatterers, separated by a distance w, and measured to a resolution $q_{\max} = 2\pi/d$. In one dimension there are only four values to describe this object: the distance between the points, the moduli of the scattering strengths of both points, and the relative phase between them. From Eqns. (4) and (7) the diffraction pattern is a fringe pattern

$$I(\mathbf{q}) \propto |\tilde{\rho}(\mathbf{q})|^2 = \left|a_1 e^{i\phi_1} + a_2 e^{i\phi 2}\, \exp(i\mathbf{w}\cdot\mathbf{q})\right|^2 = a_1^2 + a_2^2 + 2a_1a_2\cos(wq - \phi_1 + \phi_2), \quad (11)$$

which has a period $q_p = 2\pi/w$ and a contrast $(I_{\max} - I_{\min})/(I_{\max} + I_{\min}) = 2a_1a_2/(a_1^2 + a_2^2)$. The fringe pattern is also shifted from the origin by the phase difference $\phi_1 - \phi_2$. The measurements of these quantities, and the overall strength of the diffraction pattern is enough to retrieve the amplitudes of the scatterers although there is an ambiguity in the sign of the phase difference and in which point has amplitude a_1. In this example, the number of measurements matches the number of degrees of freedom in the image (ignoring the trivial differences just mentioned). Consider now an object consisting of three points in a row, each separated by w. This will give diffraction fringes spaced by $2\pi/w$ due to interference between the scattering from the extreme points, as for the two-point case, and also at half that period, π/w, from interference of scattering from neighboring points. In this case it is not possible to uniquely assign the strengths and phases of the points. For example, the two objects $O_1 = \{1, 4, 4\}$ and $O_2 = \{2, 5, 2\}$ have equal diffraction patterns, where the three values are the scattering amplitudes of the three equally-spaced points (see Fig.4 (a)). In fact, the 3-point problem generally has two solutions, since $\{1, a+b, ab\}$ and $\{a, 1+ab, b\}$ generate the same pattern. Therefore, image reconstruction is not generally unique in one dimension. In two dimensions it has been shown that the solution is most likely unique, except for rare and special cases [14]. This is essentially due to the fact that for each line in a 2D image there are orthogonal line images that constrain the image values. For example 5-point 2D objects constructed from the sequences O_1 or O_2 can be distinguished from each other through the difference in strength of the diagonal fringes in the patterns as

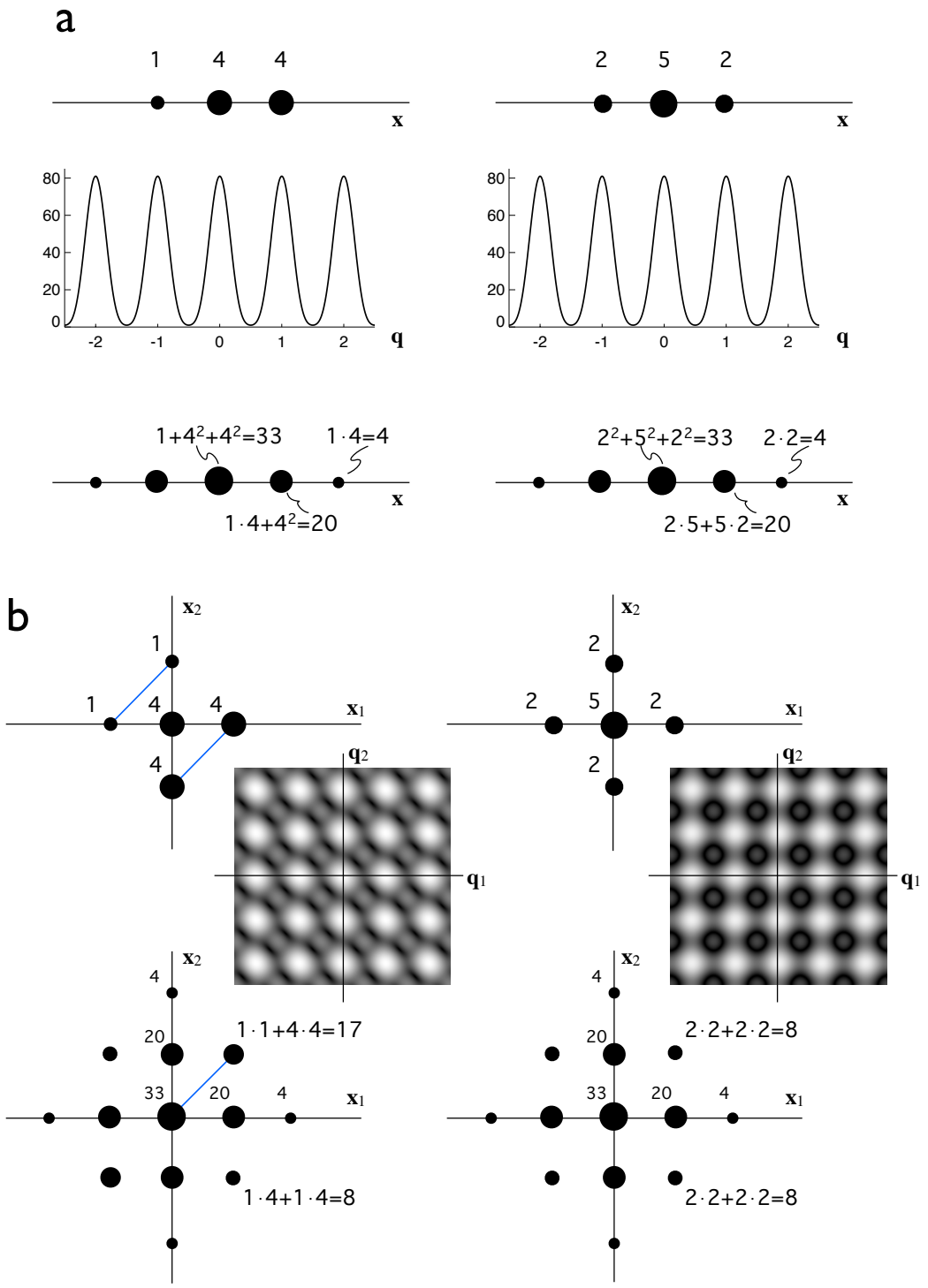


Fig. 4: *(a) Two different 3-point objects (top) have the same 1D diffraction pattern (center) and spectrum of diffraction intensities (or autocorrelation) (bottom). (b) This equivalence is broken by interference between objects in different directions, introduced in two or more dimensions.*

shown in Fig. 4 (b). While it may be possible to construct two or more 5-point 2D objects with the same pattern, these are unusual cases. Image reconstruction is even more robust in three dimensions.

The frequencies in the intensity pattern are represented by *its* Fourier spectrum $\tilde{I}(\mathbf{x})$. This real-space function is proportional to $\mathcal{F}^{-1}\{|\tilde{\rho}(\mathbf{q})|^2\} = \rho(\mathbf{x}) \otimes \rho^*(-\mathbf{x})$, the autocorrelation of the electron density (as obtained from object spatial frequencies on the Ewald sphere in the case of a single coherent diffraction pattern). We saw in the case of the two-point object that the diffraction pattern consists of just two frequencies: the zero frequency (a dc term) and one of period $2\pi/w$. $\tilde{I}(\mathbf{x})$ gives the phase and amplitude of these frequencies, as well as the negative frequency which adds no further information since $\tilde{I}(\mathbf{x})$ is Hermitian, $\tilde{I}^*(-\mathbf{x}) = \tilde{I}(\mathbf{x})$. The image retrieval problem is equivalent to finding a compact function $\rho(\mathbf{x})$ from its autocorrelation function. Each sampled $\tilde{I}(\mathbf{x})$ is dependent on all pairs of image points whose positions differ by the vector $\mathbf{x}$.

The simple examples of the point objects illustrates some of the requirements in measuring diffraction data and hints at how the missing phases are encoded in the diffraction pattern. An object of maximum width w will give rise to fringes of finest period $2\pi/w$ in the diffraction pattern. The diffraction pattern is therefore band-limited, and from Shannon's sampling theorem the complete intensity field is determined from greater than two equally-spaced samples for every $2\pi/w$ period: $\Delta q > \Delta q_S = \pi/w$ (where we refer to Δq_S as the Shannon sampling interval). Collecting samples at finer intervals than π/w does not provide any more information (although in practice there are reasons to do so, discussed below). For a resolution $q_{\max} = 2\pi/d$, there are $N_S = 2q_{\max}/\Delta q_S = 4w/d$ Shannon samples (if the diffraction pattern is measured from $-q_{\max}$ to $+q_{\max}$). The Shannon sampling interval in real space is $\Delta x_s = d/2$, and so the minimum real-space field of view that can be retrieved from adequately-sampled diffraction data is $N_S\,\Delta x_s = 2w$, twice the width of the compact object. The diffraction pattern from a general 1D complex-valued object of width w will consist of $N_S = 4w/d$ independent measurements, for a resolution of d. The number of degrees of freedom in the complex-valued image is $2w/\Delta x_s$ (a real and imaginary value per each of the $w/\Delta x_s$ real-space samples). This is exactly equal to the number of Shannon samples in the pattern, which is a necessary condition to retrieve a unique image (although, as demonstrated above, not a sufficient condition). In 2D, an object of square support of width w gives rise to N_S^2 independent measurements in its diffraction pattern, needed to recover $2(w/\Delta x_s)^2 = N_S^2/2$ image coefficients. In 3D there are four times as many independent measurements as unknowns. Another way of viewing this is that the region in real space that is constrained becomes a larger fraction of the sampled space as the number of dimensions increases: outside the support we know that the image amplitude is zero.

The ratio Ω, of the number of independent measurements to the number of independent image coefficients, appears to be a good indicator of the ability to reconstruct the image [15]. The shape of the object's support can increase this overdetermination ratio, with non-convex shapes and those without centrosymmetry leading to higher values. For example a triangle support gives $\Omega = 3$ in 2D, compared with $\Omega = 2$ for a square. Objects with well-separated components are generally easier to reconstruct. A 2D object support consisting of two squares of width w with their center points separated by $> 2w$ has $\Omega = 3$ since the autocorrelation support consists of three squares of width $2w$. The ratio Ω is increased further when when one of the components is smaller than the other, approaching the case of Fourier transform holography as one of the components approaches a delta function (see [16]). The autocorrelation vectors of magnitude greater than the width of either of the components must be due to pairs of image points in separated components, which further constrains the reconstruction.

Note that a crystal consisting of identical objects arranged in a cubic lattice of width w gives rise to Bragg peaks separated by $\Delta q_B = 2\pi/w$, which is twice the Shannon sampling rate Δq_S (the underlying molecular transform is under sampled in each dimension by a factor of 2 relative to the Shannon rate). In this case the number of unknowns to be retrieved is twice the number of independent measurements in 3D, $\Omega = 1/2$, which explains why solving the phase problem for crystals is more difficult than for non-periodic objects. Non-crystallographic symmetry (where two or more copies of a molecule occur in the crystal which are not related by the symmetry of the lattice) and changing the unit cell parameters by swelling are two methods to increase the number of independent measurements to improve phasing efforts. In many cases the molecule does not fill the entire volume of the unit cell and the smooth solvent region reduces the number of coefficients to recover. In large complexes this solvent volume can reach as high as 80%, which would increase Ω from $1/2$ to 1.25 (if the support of the molecule was known). In other instances, some proportion of the molecular structure may be known (e.g. an unknown drug fragment bound to a known target), which may significantly change the balance of known parameters to retrieved phases.

3.2 Iterative Phase Retrieval Algorithms

Reconstructing an image from a coherent diffraction pattern of an isolated object requires retrieving the diffraction phases, utilizing the constraint that the real-space image $\rho(\mathbf{x})$ is zero outside its support, plus any other constraints that can be applied. Iterative algorithms successively apply operators with the aim to converge to the image which is contained within the support and whose Fourier spectrum matches the measurement.

The simplest iterative scheme is Fienup's Error Reduction (ER) algorithm [8], which begins from the square root of the measured diffraction intensities and a random guess of the phases: $\tilde{\rho}'_1(\mathbf{q}) = \sqrt{I}(\mathbf{q}) \exp\{i\phi(\mathbf{q})\}$. The image $\rho'_1(\mathbf{x})$ is then formed by an inverse Fourier transformation. This image will not be zero everywhere outside the actual support S of the object. The estimate $\rho_2(\mathbf{x})$ is obtained by setting $\rho'_1(\mathbf{x})$ to zero outside S. This is transformed to $\tilde{\rho}_2(\mathbf{q})$ which will no longer be in agreement with the measured Fourier amplitudes. An update is formed by setting $\tilde{\rho}'_2(\mathbf{q}) = \sqrt{I}(\mathbf{q}) \exp\{i\phi_2(\mathbf{q})\}$, where $\phi_2(\mathbf{q})$ are the phases from $\tilde{\rho}_2(\mathbf{q})$. This procedure is then iterated, ideally until both sets of constraints are satisfied.

The image can be represented by a vector in an N-dimensional vector space, where N is the number of pixels (or voxels) in the image. The complex amplitude at each pixel gives the value of the vector along each corresponding dimension. In this representation the ER algorithm can be written as

$$\rho_{n+1}(\mathbf{x}) = P_S \, P_M \, \rho_n(\mathbf{x}) \tag{12}$$

where the projection operators are given by

$$P_S \, \rho(\mathbf{x}) = \begin{cases} \rho(\mathbf{x}) & \text{if } \mathbf{x} \in S \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

and

$$\tilde{P}_M \, \tilde{\rho}(\mathbf{q}) = \sqrt{\frac{I(\mathbf{q})}{|\tilde{\rho}(\mathbf{q})|^2}} \, \tilde{\rho}(\mathbf{q}) \tag{14}$$

with

$$P_M = \mathcal{F}^{-1} \, \tilde{P}_M \, \mathcal{F}. \tag{15}$$

Projection operators have the property that $P^2 = P$. Error metrics describing how well the modulus and support constraints are satisfied can be expressed as Euclidean distances $\epsilon_S = \|P_S\rho - \rho\|$ and $\epsilon_M = \|P_M\rho - \rho\|$, respectively, where $\|\cdot\|$ is defined as the sum of the square moduli of the vector components. It is clear from these definitions that an iteration of $P_S\,P_M$ will always decrease the errors ϵ_S and ϵ_M, which is why Fienup called this the Error Reduction algorithm. However, this algorithm has no method to escape from local minima, which would require a step that increases these metrics. Numerous algorithms have been proposed to overcome the stagnation problem inherent in the ER algorithm, of which the most popular is the hybrid input-output (HIO) algorithm [8], which can be expressed as

$$\rho_{n+1}(\mathbf{x}) = (P_S\,P_M + (I - P_S)(I - \beta P_M))\rho_n(\mathbf{x}), \tag{16}$$

in the case when a only a support constraint is applied in real space. Inside the support, where $(I - P_S)\rho = 0$, the modulus constraint is applied as in the ER algorithm. Outside the support, instead of setting ρ to zero to exactly satisfy the support constraint, the iterate ρ_{n+1} is formed by subtracting $\beta P_M\,\rho_n$ from ρ_n, where β is a constant usually in the range 0 to 1. The inspiration for this algorithm comes from control theory, with the idea to provide a negative feedback to the operation of applying the modulus constraint. The input to P_M is compensated at those points where the support constraint is violated. This allows the algorithm to escape local minima. Often several iterations of ER are inter-dispersed between HIO steps. Other algorithms improve the convergence rate by taking bigger steps in image space or altering the search strategy, as explained and reviewed by Marchesini [17].

Iterations proceed until the error metrics converge. The final iterate is not necessarily equal to the solution, $\bar{\rho}$, which is the intersection of the two sets, $\bar{\rho} = P_S\bar{\rho} = P_M\bar{\rho}$ [18]. The solution can be found from $\bar{\rho}(\mathbf{x}) = P_M\,\rho_n(\mathbf{x})$. Due to measurement noise a single true solution cannot be distinguished from a family of images that satisfy all constraints to within the the errors. An average solution can be determined by continuing the iterations and generating solutions say every 100 iterations, or by rerunning the algorithm from random phases [19]. This average is unique, and is perhaps the best estimate of the true image that could be obtained from the data. The procedure also allows us to determine the reliability of the retrieved phases. Diffraction phases that always reconstruct to the same value will add complex amplitudes constructively, whereas the sum of amplitudes with random phases will tend to zero. By comparing the modulus of this average to the square root of the measured diffraction intensities we obtain a value at each diffraction pixel that is < 1 and only equal to 1 when phases are exactly consistent. This value tends to decrease with increasing resolution due to the fact that diffraction amplitudes decrease with resolution and are more influenced by noise. This ratio indicates how well the spatial frequencies of the image are represented in the phasing process and is accordingly referred to as the phase retrieval transfer function (PRTF). The average solution can be thought of as the true solution imaged through an optical system with a transfer function given by the PRTF. When performing the average the constant phase term must be normalized to the same value for each solution, otherwise the PRTF will be < 1 for the zero frequency (representing an attenuating optical system). Other low-order aberrations including wavefront tilt (displacement of the image), defocus, and Seidel aberrations such as astigmatism and coma, can also be removed from each solution before determining $\langle\bar{\rho}\rangle$. Since these types of aberrations only cause blurring or shifting of the image, they are not fully constrained by a loose support, unless a positivity constraint is applied (the out of focus image of a positive-density object will have negative amplitudes) [20].

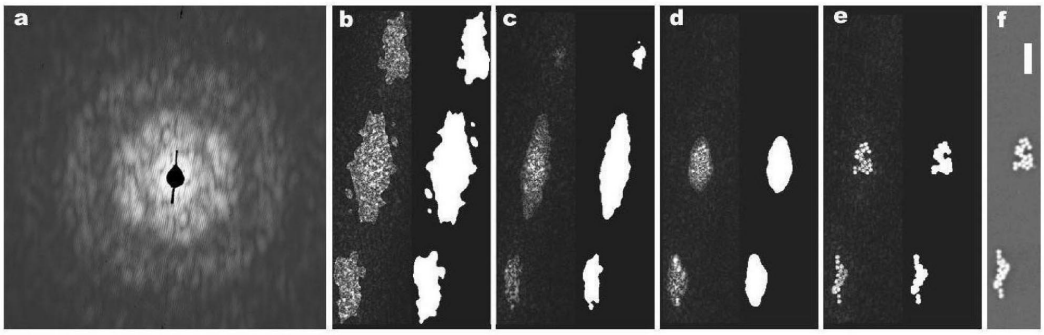


Fig. 5: *Image reconstruction from the coherent diffraction pattern in (a). The process starts from zero phases, giving the autocorrelation of the image, in (b). The first support is obtained by a threshold of this image. As phase-retrieval iterations proceed the support is re-estimated, as shown in (c) to (e). An SEM image of the object is in (f). From [5].*

When imaging at hard X-ray wavelengths away from atomic resonances, or imaging binary objects at any wavelength, a positivity constraint can be applied to the real-space image values. With soft X-rays (photon energies below about 1 keV) most interesting objects are complex valued. In this case, a successful reconstruction requires a support constraint that closely matches the actual boundary of the object, since with a loose support an out-of-focus image would be a valid solution. However, it is unlikely that the support is known to such accuracy. Marchesini noticed that the tighter the support constraint the better the reconstruction, leading him to propose the Shrinkwrap method [5]. As illustrated in Fig. 5 this adds a dynamic support constraint to an iterative transform algorithm, where the support is updated occasionally based on the current reconstruction. Image pixels below a certain threshold of a blurred version of the current reconstruction are treated as being outside the object. Starting from the support of the image autocorrelation, the support tends to gradually shrink to the boundary of the object, improving the reconstruction as it does so (which then gives an improved estimate of the support). A stopping criterion must be added otherwise the method tends to over shrink; when this happens the error ϵ_S increases abruptly, allowing this transition to be determined.
In many experimental situations, especially single-shot imaging at X-ray free-electron lasers, there are missing data at low scattering angles near the zero-order beam. Unless the sample has been designed to be mostly absorbing (as can be achieved in imaging magnetic domains in thin films, for example) the zero-order beam will be extremely intense, and is usually managed with a beam stop or a hole in the detector. If this missing region is as small as a single speckle, only the overall scattering strength of the object is lost. As this region increases more information is lost, often causing iterative algorithms to fail. When only a support constraint is applied, there will be particular modes that are neither constrained by the diffraction measurement nor the support [19]. However, it is possible to reconstruct images missing these low frequencies, just as a lens with an annular pupil will give rise to high-pass filtered images [21].

3.3 Coherence and Detector Requirements

Reconstructing images of non-periodic objects from their X-ray diffraction patterns has been found to be more problematic than experienced in computer simulations or visible light experiments. This may be due to the fact that most X-ray experiments have been carried out with

partially-coherent beams. The van Cittert-Zernike theorem states that the mutual coherence function of a beam propagating from an incoherent source is given by the Fourier transform of the source intensity distribution. Application of this theorem to an object illuminated by this partially-coherent beam shows that its far-field diffraction pattern is given by

$$I(\mathbf{q}) = |\tilde{\rho}(\mathbf{q})|^2 \otimes p(\mathbf{q}) \tag{17}$$

where p is the spatial distribution of the source, such that its angular extent is given by $p(4\pi\theta/\lambda)$. More generally, in the Gaussian-Schell model the source can described in terms of mutually incoherent modes and the diffraction pattern is then the sum of intensity patterns arising from each mode. The image reconstruction as described in the previous section assumes full coherence, with $p(\mathbf{q}) = \delta(\mathbf{q})$. As can be seen from Eqn. (17) the effect of the source is to reduce the contrast of the diffraction pattern, and in particular causing the zeroes in the intensity pattern to have positive values. This causes problems in assigning phases to the values $\sqrt{I(\mathbf{q})}$ since there will be a discontinuity from negative to positive diffraction amplitudes.

The effect of the finite width of detector pixels is also a convolution, given by Eqn. (17), where $p(\mathbf{q})$ is now the pixel sensitivity. The decrease in fringe contrast in the diffraction pattern causes a modulation of the autocorrelation by the function $\tilde{p}(\mathbf{x})$, the modulation transfer function of the detector or the mutual coherence function of the source. The width of this function (i.e. the transverse coherence length) must be significantly wider than the width of the object autocorrelation, or twice the largest diameter of the object. For example, a square-pixel detector that is fully sensitive across its area (which is approximately the case for CCD detectors) will not detect the finest fringes in the pattern when the pixel spacing matches the Shannon sampling rate. That is because a pixel integrates over half a period of these finest fringes, which will be exactly equal to the integration over the other half period in its neighboring pixel. Improved estimates of the Shannon samples are obtained by increasing the pixel density.

Often the function p can be estimated or measured so that the diffraction pattern contrast can be corrected. This deconvolution is carried out by dividing the autocorrelation $\tilde{I}(\mathbf{q})$ by the MTF $\tilde{p}(\mathbf{q})$, a high-pass filtering operation. Since this deconvolution procedure can amplify noise, following Whitehead *et al.* [22] we could propose to *convolve* the current iterate with p before comparing with the measurement to update the modulus of $\tilde{\rho}$, by replacing Eqn. (14) with

$$\tilde{P}_M\, \tilde{\rho}(\mathbf{q}) = \sqrt{\frac{I(\mathbf{q})}{|\tilde{\rho}(\mathbf{q})|^2 \otimes p(\mathbf{q})}}\, \tilde{\rho}(\mathbf{q}). \tag{18}$$

In this way the moduli are updated by first computing the partially-coherent diffraction pattern that would arise from the current estimate of $\rho(\mathbf{x})$, and comparing this with the measurement $I(\mathbf{q})$. Equation (18) is not a projection operator, since repeated application may continue to change ρ. The modal method for reconstruction of images from partially-coherent diffraction patterns derived by Whitehead *et al.* [22] has nevertheless been found to substantially improve the success of image reconstructions using experimental data. When the width of p is small enough that the highest period fringes discerned in the pattern are due to the most distant inter-object separation, then even rough estimates of p lead to improved reconstructions.

The assumption of quasi-monochromaticity inherent in the discussion up until now requires a bandwidth spread $\Delta\lambda/\lambda < 1/N_S$. This condition can be relaxed if the spectrum is known such that the polychromatic diffraction pattern due to the current estimate of $\rho(\mathbf{x})$ can be calculated [23], in a similar treatment to partial spatial coherence. Additionally, the framework of partial coherence can be applied to changes in the sample itself, either due to X-ray induced damage [24] (and see below) or sample motion [25].

4 Outrunning Radiation Damage

4.1 Diffraction Before Destruction and Time-Delay Holography

As mentioned in the Introduction, recording diffraction from non-crystalline material requires a vastly greater exposure than required for crystalline material. The limitation to acquiring high-resolution diffraction information is due to radiation damage: the very radiation used to determine structure impairs that structure. The dose (energy per unit mass) that cryogenically cooled biological material can withstand is about 30 MGy before changes are observed in the molecular constituents of the sample. By using pulses shorter than the timescale of radiation damage processes, we can literally outrun damage and increase the dose by many orders of magnitude [11]. With short-pulse X-ray FEL radiation at high irradiance (10^{18} to 10^{21} $\mathrm{W/cm^2}$ or doses up to 100 GGy at 8 keV photon energy) the high degree of ionization of every atom in the sample will cause an isolated object to Coulomb explode, yet the inertia of atoms provides on the order of 10 fs before atomic displacements will significantly alter the diffraction pattern at atomic resolutions. The electron density of ions is of course lower than neutral atoms, and this ionization will also give rise to modification to the pattern even before atomic motion occurs. The random occurrence and sequence of ionization will on average cause an uncorrelated (that is, q-independent—at least at low resolution) addition of diffuse diffraction and reduce the overall contrast. This method of "diffraction before destruction" was first demonstrated at the FLASH FEL at soft X-ray wavelengths [26], and has now been verified to hold at atomic resolution [27], and the processes of radiation damage over the timescale of the pulse is an active area of experimental and theoretical research.

Since initial FEL experiments were limited to long wavelengths (and hence low spatial resolution), essentially no damage could be observed over the limited duration of the pulse [26]. The effects of the X-ray pulse on the material could only be discerned by tracking the explosion at tens of nanometre length scales taking place on picosecond timescales. This was carried out by an interesting method dubbed time-delay holography [28]. After interacting with the sample, the pulse is reflected by a normal-incidence multilayer mirror back onto the sample to diffract again. The initial (prompt) diffraction is also reflected in the mirror. This and the delayed diffraction propagate together to a diffraction camera located in a backscattering geometry (the incident beam passes through a hole in this detector on the way to the sample). The time delay between the pulse passing through the sample for the first and second times is given by $2\ell/c$ where ℓ is the distance from sample to mirror and c the speed of light. In experiments this was varied from 350 fs to several picoseconds. The two diffraction patterns interfere at the detector since there is zero path difference between the light that diffracts first from the sample then reflects and the light that reflects first from the mirror then diffracts. Unfolding the effect of the mirror shows we are diffracting from two objects longitudinally displaced a distance 2ℓ from each other. Two point sources in this configuration give rise not to straight fringes, but circular fringes with a spacing that decreases with the square of scattering angle (the center of the circles is on the line passing through both objects, i.e. normal to the mirror). Such a pattern is convolved with each scattering point in the sample. The fringes precisely encode the time delay, and the prompt diffraction (from the known undamaged object) acts as a holographic reference to help retrieve an image of the exploding object. A shift of the fringes gives an interferometric measurement of the change in refractive index of the object, which was used to determine that the expansion of a latex microsphere test object was less than 6 nm in 350 fs. The idea for this experiment was inspired by an exhibit of Newton's dusty mirror at a science museum where

circular fringes could be observed (in visible light) from spores dusting the front surface of a back-silvered mirror. Newton first observed and described interference via this arrangement, although Thomas Young was the first to interpret the observations correctly. Our method of reflecting the same pulse onto the sample to examine it at a precise later time was also used to confirm predictions that a sacrificial layer around an object (a tamper) delays the explosion of that object [29].

4.2 Diffraction Termination

Molecular dynamics calculations can give detailed simulations of the X-ray induced explosion that take into account the initial molecular structure of samples. These calculations are computationally expensive, and so are limited to small samples such as single macromolecules. A continuum approach applies methods developed to model dense plasmas such as in stellar interiors (somewhat larger than a molecule), but makes the assumption that in small regions the dynamics of the sample are isotropic. A code called Cretin was used to simulate the structural changes at atomic resolution in small protein crystals of about 1 μm diameter, embedded in a water tamper [12]. The calculations make discrete time steps in which the atomic populations and electron and photon distributions are calculated utilizing known transition rates and opacities. These provide electron and ion temperatures, ionization states and ion collisions, from which mean atomic displacements are computed, as shown in Fig. 6. The RMS displacements are obtained from a diffusion equation where the diffusion coefficient is calculated from the ion collision frequencies and temperatures. As seen in the figure the RMS atomic displacement σ increases approximately as $t^{3/2}$ over the time t of a constant-irradiance pulse. The explosion occurs faster for higher pulse irradiance, with σ approximately proportional to the square root of pulse irradiance (photons/unit area/unit time) for the range of irradiance considered here. At 8 keV photon energies, this dependence continues beyond an irradiance of $10^{21}\,\mathrm{W/cm^2}$ which corresponds to focusing 10^{13} photons in 10 fs to a spot of 0.4 μm diameter. At this photon energy and irradiance, an RMS displacement of 1 Å is reached in about 10 fs.

Each X-ray induced explosion of an identical object will be different due to the random sequence of ionization and atomic displacements. The diffraction signal obtained by averaging over many instances of these explosions can be obtained through the derivation, found in several text books (see e.g. [30]), for describing small random displacements in crystals due to thermal motion. It is interesting to note that this common derivation makes no explicit assumption as to whether the almost-identical objects are arranged and exposed together (as in a crystal) or exposed in a serial fashion and diffraction intensities then summed. This is the case if there is no spatial correlation to the displacements and hence no correlation of these displacements from unit cell to unit cell (if considering a crystal), and that the RMS displacement within a single object is the same as the RMS displacement in the whole crystal. We assume that diffraction patterns are properly oriented prior to averaging, by the methods of Sec. 5. The effect of isotropic displacements is to multiply the diffraction signal at a momentum transfer q by a factor of $\exp(-q^2\sigma^2)$. Here σ is the RMS of the component of the displacement in the direction of $\mathbf{q}$, equal to the 1D component of the mean displacements for isotropic displacements, as plotted in Fig. 6. In our accelerating explosion, σ increases during the exposure. The instantaneous diffraction signal at a resolution d is reduced by $1/e$ when σ reaches a value of $d/(2\pi)$. Displacements of only 0.5 Å terminate the accumulation of intensities at 3 Å resolution. The summed diffraction signal from recorded patterns is the pulse-integrated signal. Following

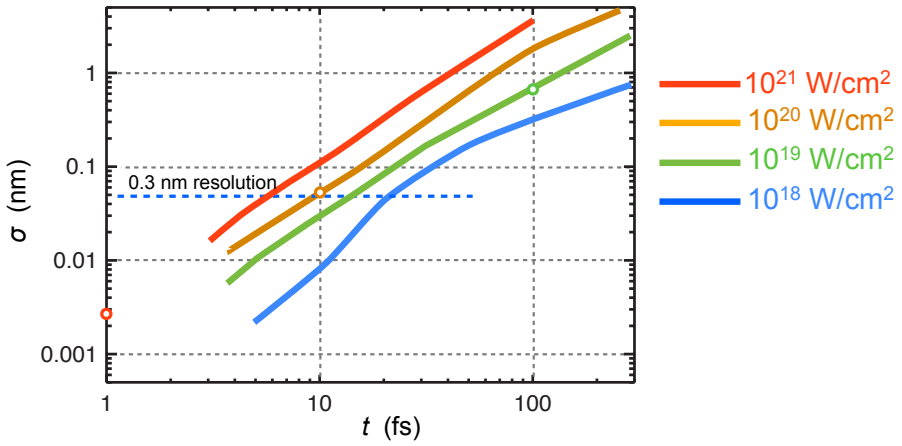


Fig. 6: *Plot of the RMS isotropic atomic displacement σ as a function of time t during pulses of constant irradiance as shown in the legend, for a photon energy of 8 keV. The open circles indicate the time required to reach a fluence of 10^6 J/cm^2 as in the simulation of Fig. 3, showing that higher pulse power (photons per unit time) gives less damage for a given total scattering signal. Lower resolution can tolerate longer (and therefore less intense) pulses. From [12].*

Eqn. (7) this diffraction signal for a constant irradiance pulse of duration T is given by

$$I(\mathbf{q}) = I_0\, T\, \Omega_p\, P\, r_e^2\, \left\{ |\tilde{\rho}(\mathbf{q})|^2\, g(q;T) + N_{\rm atom} |f|^2 (1 - g(q;T)) \right\}, \tag{19}$$

$$g(q;T) = \frac{1}{T} \int_0^T \exp\{-q^2 \sigma^2(t)\}\, dt. \tag{20}$$

Here we have omitted another multiplicative Wilson factor of $\exp(-Bq^2/8\pi^2)$ to describe the structure variation inherent in the unexposed samples, and note that the total incident fluence is given by $I_0\, T$ (in units of photons per unit area).
Equation (19) consists of two terms. The first is the undamaged diffraction pattern $|\tilde{\rho}(\mathbf{q})|^2$ modified by the dynamic disorder factor $g(q;T)$, a function that decreases monotonically with increasing q. The second term is a slowly-varying background of scattered counts that is proportional to the overall scattering strength of the sample and which increases monotonically with increasing q. This term is 0 at $q = 0$. Thus, in general, the disordering takes scattered counts out of the information-containing Shannon samples into a background that diminishes the contrast. This occurs first at high q and works its way towards lower resolution with time. Applying the empirical time dependence observed in Fig. 6 of $\sigma(t) = \sigma_T \cdot (t/T)^{3/2}$, with σ_T the RMS displacement at the end of the pulse, we find

$$g(q;T) = \int_0^1 \exp(-q^2 \sigma_T^2 t'^3) dt' = \frac{\Gamma(4/3)}{(q^2\sigma_T^2)^{1/3}} - \frac{1}{3} E_{2/3}(q^2\sigma_T^2) \tag{21}$$

where $E_{2/3}(x)$ is the exponential integral function of order $2/3$, and which tends to zero as the argument x increases, and $\Gamma(4/3) = 0.89$. That is, for high resolution $g(q;T) \approx (q\,\sigma_T)^{-2/3}$.
The behavior of g can perhaps be better understood from Eqn. (20). At early times in the pulse $\sigma(t)$ is small and the instantaneous disorder factor (the integrand) is close to unity and the diffraction pattern, proportional to $|\tilde{\rho}(\mathbf{q})|^2$, continues to accumulate. When σ reaches a value of $1/q$ this accumulation stops and the background then accumulates. This happens at a

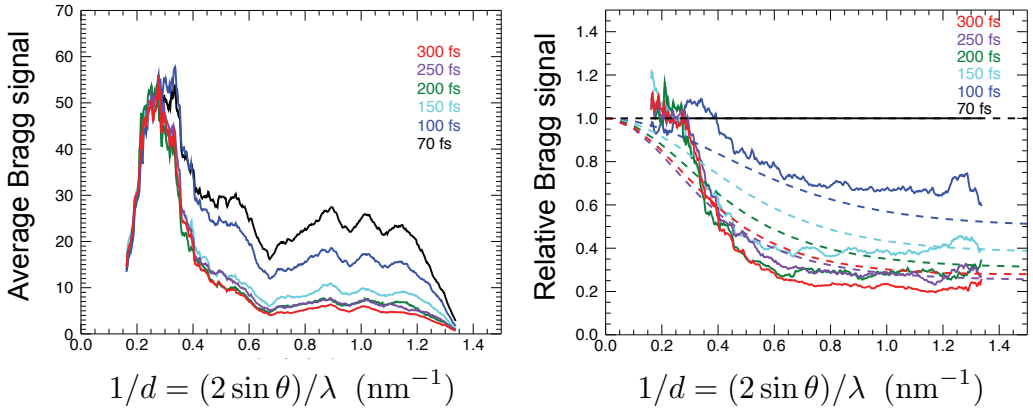


Fig. 7: *(left) Average Bragg signal of Photosystem I micro crystals as a function of resolution for different pulse durations for pulse irradiance up to* $10^{17}\,\mathrm{W/cm^2}$ *at a photon energy of 2 keV. (right) The ratio of the average Bragg signal to that at the shortest pulse duration, and comparison to ratios of the dynamic disorder factor* $g(q;T)$ *obtained from continuum modeling of the evolution of the RMS atomic displacements as shown in Fig. 6. From [12].*

time t_{off} such that $1/q = \sigma_T(t_{\mathrm{off}}/T)^{3/2}$, or $t_{\mathrm{off}} = (q\,\sigma_T)^{-2/3}\,T$. The proportion of the pulse that contributes to the cooperative diffraction is given by t_{off}/T, in agreement with the limiting value of the integration of Eqn. (21). Pulses longer than t_{off} do not contribute any more information at resolutions of q or better, and simply accumulate background for the rest of the pulse.

Our analysis suggests that the source metric to maximize signal acquisition is not the total pulse fluence, but the pulse irradiance (or source power, considering that the spot size is essentially dictated by focusing optics). If we simply increase the pulse energy (number of photons) by proportionally increasing the pulse length, then we only increase the background without improving the signal. Increasing the irradiance increases the rate of signal photons arriving on the detector but does shorten the time t_{off} that these photons accumulate. However, given the linear dependence of σ_T^2 on pulse irradiance observed in Fig. 6 we find that the total signal scales as $I_0\,t_{\mathrm{off}} \propto I_0^{2/3}$. However, if irradiance is scaled higher than $10^{21}\,\mathrm{W/cm^2}$ (at 8 keV photon energy) we can expect almost every electron to be stripped from every atom in the sample. In this case the turn-off time of diffraction will not be necessarily limited by atomic motion, but the variability in the atomic scattering factors at undisplaced atomic positions. This has no spatial correlation, and leads to a diffuse background that is independent of q and a corresponding uniform decrease in the signal. In this case, our q-independent turn-off time will depend on the atomic relaxation processes, which are not much longer than several femtoseconds.

4.3 Bragg termination

Among the first experiments to be carried out at the LCLS, the world's first hard X-ray FEL [1], was to measure the effect of pulse duration and fluence on coherent diffraction patterns. The samples chosen for confirming "diffraction before destruction" and mapping out the behavior of diffraction termination were protein nanocrystals flowing in a water jet. These are biological material and are confined in a water medium, as per the system modeled in Fig. 6. The crystallinity gives a great advantage that the diffraction signal $|\tilde{\rho}(\mathbf{q})|^2$ is confined to Bragg

peaks, whereas the background due to the explosive disordering varies very slowly across the detector. The presence of discernible Bragg peaks indicates that structure at a particular length scale persisted for some fraction of the pulse. Indeed, the measurement of the Bragg intensities as a function of pulse duration allowed us to determine this fraction, and hence quantify $t_{\rm off}$ and σ_T. For even small protein nanocrystals, and pulses not very much longer than $t_{\rm off}$, the disorder background is much weaker than the Bragg peaks. Therefore, for nanocrystallography it is not necessary to limit pulse durations to $t_{\rm off}$. The Bragg peaks terminate when the disorder reaches a level of $\sigma = d/(2\pi)$ for a resolution length d. The effect of the explosion is a filtering of the Bragg values by the dynamic disorder factor $g(q;T)$, which can be corrected by dividing the Bragg signal by an estimate of this function based on modeling of the explosion or from measurements at several pulse durations. In addition, the method of nanocrystallography opens up protein structure determination to samples that cannot be grown into crystals of sufficient size or quality for standard synchrotron radiation measurements or are particularly sensitive to radiation damage, making this a very attractive and important method for biological structure determination [27].

Figure 7 shows measurements of the termination of Bragg peaks as a function of resolution for different pulse durations, carried out at a photon energy of 2 keV (6 Å wavelength) on nano and microcrystals of Photosystem I. This is a large membrane complex involved in photosynthesis, and chosen for these experiments primarily because of its large unit cell and because it has proven to be a very difficult sample for conventional synchrotron measurements which required crystals several hundred microns large. As can be seen from Fig. 7 the diffraction efficiency is very much lower at high resolution for longer pulses compared with the shorter pulses, simply because only the first part of the pulse is contributing to the diffraction. Even so, the diffraction patterns with 300-fs pulses give Bragg peaks, at 1-nm resolution, that are easily measured and well above background. The model predictions contain no fitted parameters, and the agreement with the experimental results gives confidence of the predictions for atomic-resolution imaging.

5 Assembly of Diffraction Data from Identical Particles

As we have seen, the intense and extremely brief pulses from X-ray FELs solve the problem of radiation damage, but a single object only survives a single pulse. In the scattering geometries discussed here, this limits information collected in a single pulse to a two-dimensional slice (on the Ewald sphere) through the three-dimensional diffraction intensities $|\tilde{\rho}(\mathbf{x})|^2$. The acquisition of full three-dimensional information requires combining data from many copies of identical objects in different orientations. Summing data from many objects also increases the signal beyond that available in a single-shot pattern, which is limited by the number of incident photons per pulse as estimated by Eqn. 9. There are several ways in principle to acquire data from multiple particles so that scattering information can be combined into a single 3D dataset. We consider here methods in which the orientation of the particle to the frame of the laboratory is not known *a priori*, referred to by Elser as *cryptotomography* [31].

5.1 Serial Nanocrystallography

The most familiar and most successful method for increasing the diffraction signal of macromolecules and their assemblies is by forming crystals of these samples. Every copy in the crystal is oriented with respect to all others, and these are arranged in a 3D lattice which gives

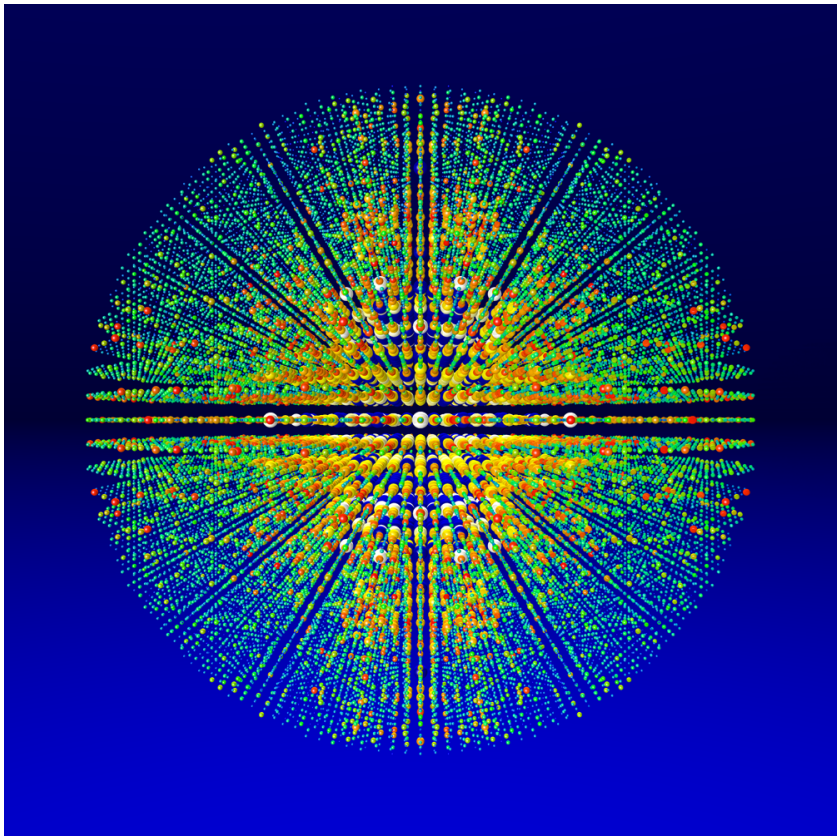


Fig. 8: *The assembled diffraction data from 56,000 nanocrystal diffraction patterns of Photosystem I collected at LCLS at a photon energy of 2 keV. Figure by Thomas White, CFEL DESY.*

constructive addition of the scattering in the particular locations of Bragg peaks. Since the real-space electron density of the crystal is a convolution of the repeated motif with a lattice, the Fourier representation $\tilde{\rho}(\mathbf{q})$ is the transform of the motif modulated by the Fourier transform of the lattice. The total integrated signal in each Bragg peak is N_{mol} times greater than that of the single particle, where N_{mol} is the total number of repeats in the crystal, and furthermore this signal is concentrated into a Fourier-space volume much smaller than the Shannon volume. For short X-ray FEL pulses, the crystalline diffraction (from structure that is correlated from unit cell to unit cell) will always give much higher counts per pixel than the disordered background. This background (which is also proportional to the total number of atoms in the crystal—see Eqn. (19)) uniformly fills the detector area between Bragg peaks whereas the peaks themselves are concentrated into approximately $1/N_{\mathrm{mol}}^{2/3}$ of the area between peaks. That is, a $10 \times 10 \times 10$ crystal will have peaks about 100 times above background when $g = 1/2$. The well-known disadvantage of crystals, mentioned in Sec. 3.1, is that the Bragg-peak spacing Δq_B is twice the Shannon spacing Δq_S for an object with the same width as the real-space lattice period, leading to fewer independent measurements than image coefficients.

In serial crystallography [27], crystals of a few micrometer or smaller are delivered to the X-ray beam in a jet of a water. A crystal is hit by chance and is situated in a random and unknown orientation. The determination of this orientation can be readily made from the pattern itself by indexing the Bragg peaks [32]. Several so-called auto-indexing programs have been developed for crystallography which search for a repeating lattice in the measured diffraction

pattern, knowing the mapping of that pattern onto the Ewald sphere. In the simple example of a monoclinic crystal with different unit cell spacings in each dimension, the reciprocal lattice basis vectors can be identified quite readily from the observed reciprocal lattice spacings and arrangement. Each Bragg peak in the pattern can thereby be indexed by its 3D Miller index $\mathbf{h}$ and thereby properly accumulated in a running sum of $|\tilde{\rho}(\mathbf{h})|^2$. In this way the diffraction data are built up, averaging over all crystal shapes and sizes, just as in the case for X-ray powder diffraction. The difference in this case is that the summation is done in three dimensions instead of the single magnitude of q, giving us complete 3D information as if collecting data from a rotating single undamaged crystal. An example of this 3D assembly is shown in Fig. 8.

5.2 Single-Particle Diffraction

Single-particle X-ray diffraction aims to carry out the same acquisition and analysis steps as for serial nanocrystallography, without the advantage of a lattice to guide us to the orientation, and without the addition of scattering from many unit cells that is concentrated into narrow Bragg peaks. It is the special case of the single-unit cell crystal. It is of particular interest because it could be applied to any series of identical objects that cannot be crystallized and allows measurement of all independent Shannon samples for the application of the phase-retrieval methods described in Sec. 3.2. The orientation analysis must be carried out on patterns with much lower signal than a single photon per pixel, as is the case for pulse energies that can be generated at X-ray FELs, such as shown in Fig. 3.

One method for determining the relative orientation of patterns is by searching for common arcs of intersection. Since every diffraction pattern samples information on an Ewald sphere that passes through the origin of Fourier space, any two patterns will always intersect along an arc that also passes through the origin. If $\rho(\mathbf{x})$ is real valued, as it is away from atomic resonances and for hard X-ray wavelengths, then $|\tilde{\rho}(\mathbf{q})|^2$ is centrosymmetric, and there will be another arc of intersection when forming the centrosymmetric inversion of one of the patterns. (Note that the diffraction pattern itself will not have centrosymmetry because the Ewald sphere is not centrosymmetric.) These two intersections allow the identification of the relative 3D orientation of the patterns. The identification of the two arcs does work even if the object density is not strictly real-valued, as shown in Fig. 9. Identifying the arcs of intersection requires a minimum signal to noise level in the patterns, which could be built up by first classifying patterns into groups of similar enough orientation that they can be summed together without knowing their relationship to other patterns. This is the approach common in the similar method in cryo-electron microscopy, where noisy real-space images are first classified into groups of similar images and then the class averages are oriented. It has been found that when working directly with diffraction amplitudes, however, that the process of correlation and orientation can be merged into one process. Recent computational experiments [34] suggest that the correlations between pairs of patterns can be made on arcs, from which a consistent set of orientations for all patterns can then be found. After arbitrarily fixing the orientation of one pattern, $N-1$ estimates of the orientation of each of the remaining $N-1$ patterns can be obtained. This procedure makes use of all correlations between pairs of patterns.

Other analysis methods for cryptotomography include the expansion-maximization-compression (EMC) framework and topological mapping [31]. The EMC method seeks to build up a model $W(\mathbf{q})$ of $|\tilde{\rho}(\mathbf{q})|^2$ by placing the two dimensional diffraction pattern measurements over a distribution of their most likely orientations. The method is iterative, starting from unknown orientation assignments. In each iteration, the model W is expanded into diffraction patterns that would

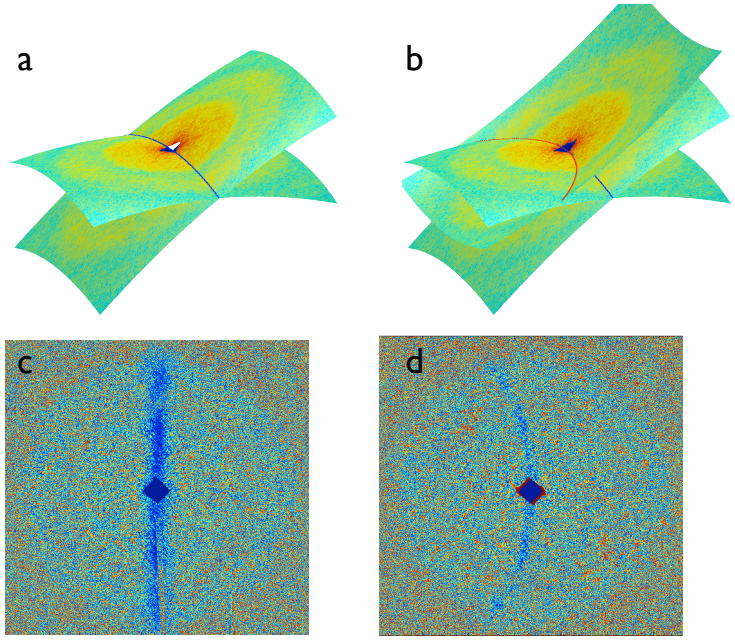


Fig. 9: *Identification of common arcs of intersection in experimental X-ray diffraction patterns taken from the same 3D nano object at two orientations. The object and diffraction data are the same as in Fig. 2. (c) shows the difference of diffraction intensities,* $I_1(\mathbf{q}) - I_2(\mathbf{q})$ *and (d) shows* $I_1(\mathbf{q}) - I_2(-\mathbf{q})$*. Figure by Gösta Huldt, Uppsala University.*

arise from such an object over a finite set of predetermined uniformly distributed orientations. The maximum likelihood step follows in which the model is changed to best agree with each of the model tomograms. In the final compression step the model tomograms are made consistent with a single updated model $W'(\mathbf{q})$. This method has successfully been applied to reconstruct the relatively simple 3D diffraction data of ellipsoidal iron oxide nano-particles measured at the FLASH FEL [33]. Elser has determined the lowest signal limits required to carry out such a procedure and finds that objects can be recovered at counts considerably less than a single photon per pixel per pattern [31]. A method that appears conceptually different (but is related [35]) is a topographic mapping approach. Utilizing the concept described in Sec. 3.2 of a diffraction pattern as a vector in a finite-dimensional vector space, it is noted that the diffraction patterns (or images) obtained at different orientations must map out a continuous 3D manifold in that higher-dimensional space. Using generative topographic mapping methods, in which this manifold is mapped out from noisy measurements, Fung *et al.* [36] recovered the structure of a molecule from simulated patterns with less than 0.04 counts per pixel.

6 Outlook

The LCLS, the first hard-X-ray FEL in the world, is now operational and many groups are carrying out new explorations into the possibilities and methodologies of imaging with brilliant coherent X-ray pulses. There is still much development needed before single-particle structure determination can be carried out routinely, but we now have strong foundations for these endeavors, as described here. All the steps for structure determination have been elucidated

theoretically with experimental confirmation. The compelling idea that the extreme irradiance of X-ray FEL pulses would allow us to obtain more information from a sample than possible at conventional sources, by virtue of outrunning the explosion caused by that high irradiance, has been examined in detail over a range of photon wavelengths, pulse durations, and fluences. These measurements give confidence that "diffraction before destruction" is valid to atomic resolution. More studies are required to determine if the dynamics of the explosion are dependent on the initial structure, especially in molecules with heavy atoms which could produce local centers of high charge that could consistently repel their surrounding atoms. However, even if there are correlated motions that occur during the explosion, the rapid destruction of cooperative diffraction by an RMS displacement of $d/(2\pi)$ due to motion of the majority of the atoms (which are of roughly equal mass) means that errors at a particular resolution will be small unless the correlated motion can be 2π times faster than the uncorrelated motion. We have found that nanocrystal diffraction is particularly impervious to the effect of the X-ray induced explosion since the diffraction from the periodic undamaged component of the structure is easily discriminated from the scattering of the disordered component. Our model of diffraction termination shows that the metric to optimize sources for imaging is pulse power. Today, the LCLS produces 50 GW X-ray pulses and designs have been proposed that could produce pulses beyond 1 TW. [37, 38].

It is straightforward to calculate the diffraction pattern of a molecular assembly (from Eqn. (7)), showing that signals at atomic resolution will invariably be at extremely low photon counts. The theoretical tools have been developed to combine data from serial measurements of identical objects to build up the three-dimensional Fourier spectrum of the average object. The experimental challenges lay in being able to actually record such weak diffraction patterns without the introduction of extraneous noise sources. Is it really possible to illuminate a macromolecular object with 10^{13} photons and yet ensure that the single photons measured in the detector pixels have scattered from that object, and are not due to scattering from optical elements, gas molecules, or other means? Once the data assembly has been accomplished, the reconstruction of a three-dimensional object is ensured by the phase retrieval algorithms that have been adequately demonstrated in many contexts. Real experimental effects such as wavelength bandwidth, spatial coherence, and detector response can all be characterized and accounted for. The phase retrieval of diffraction data sample at or beyond the Shannon rate gives an overdetermined dataset in two and three dimensions, and *ab initio* image retrieval is possible. That is, the phase problem is more easily solved than in the case of crystals, where over 50 years of insights and breakthroughs can additionally be drawn upon or adapted. For example an extension of the method of anomalous diffraction to extreme irradiance has been proposed [39].

These ideas and underpinnings have already given us a new route to obtain high-resolution room-temperature structural information from protein nanocrystals that are too small to be analyzed at conventional sources. The short pulse duration from X-ray FELs gives inherently high temporal resolution, which can be exploited in pump-probe experiments where a sample is stimulated by a short laser pulse at a precise time before the arrival of the X-ray pulse. The femtosecond pulses are over 1000 times shorter than synchrotron pulses, giving access to a full exploration of the motions involved in chemical reactions, molecular vibrations, and laser-matter interactions. The serial method can be applied to measurement of irreversible reactions. Since the sample is destroyed by the pulse anyway, there is no requirement to bring it back to the ground state as is the case for stroboscopic measurements. These experimental techniques will continue to evolve as the availability of X-ray FELs increases, providing a very bright future for X-ray imaging, in more ways than one.

References

[1] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F. J. Decker, Y. Ding, D. Dowell, *et al.*, Nat. Photon. **4**, 641 (2010).

[2] D. Fausti, R. I. Tobey, N. Dean, S. Kaiser, A. Dienst, M. C. Hoffmann, S. Pyon, T. Takayama, H. Takagi, and A. Cavalleri, Science **331**, 189 (2011).

[3] X. Yang, Z. Ren, J. Kuk, and K. Moffat, Nature **479**, 428 (2011).

[4] D. Sayre, H. N. Chapman, and J. Miao, Acta Cryst. A **54**, 232 (1998).

[5] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, Phys. Rev. B **68**, 140101 (2003).

[6] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, *et al.*, J. Opt. Soc. Am. A **23**, 1179 (2006).

[7] R. W. Gerchberg and O. Saxton, Optik **35**, 237 (1972).

[8] J. R. Fienup, Opt. Lett. **3**, 27 (1978).

[9] J. M. Holton, J. Synch. Rad. **16**, 133 (2009).

[10] M. R. Howells, T. Beetz, H. N. Chapman, C. Cui, J. M. Holton, C. J. Jacobsen, J. Kirz, E. Lima, S. Marchesini, H. Miao, D. Sayre, D. A. Shapiro, *et al.*, J. Elec. Spec. Rel. Phenom. **170**, 4 (2009),.

[11] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, Nature **406**, 753 (2000).

[12] A. Barty, C. Caleman, A. Aquila, N. Timneanu, L. Lomb, T. A. White, J. Andreasson, D. Arnlund, S. Bajt, T. R. M. Barends, M. Barthelmess, M. J. Bogan, *et al.*, Nat. Photon. **6**, 35 (2012).

[13] D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A. M. Neimann, and D. Sayre, Proc. Nat. Acad. Sci. **102**, 15343 (2005).

[14] R. H. T. Bates, Optik **61**, 247 (1982).

[15] V. Elser and R. P. Millane, Acta Cryst. A **64**, 273 (2008).

[16] I. McNulty, J. Kirz, C. Jacobsen, E. H. Anderson, M. R. Howells, and D. P. Kern, Science **256**, 1009 (1992).

[17] S. Marchesini, Rev. Sci. Instrum. **78**, 011301 (2007).

[18] V. Elser, J. Opt. Soc. Am. A **20**, 40 (2003).

[19] P. Thibault, V. Elser, C. Jacobsen, D. Shapiro, and D. Sayre, Acta Cryst. A **62**, 248 (2006).

[20] S. Marchesini, H. N. Chapman, A. Barty, C. Cui, M. R. Howells, J. C. H. Spence, U. Weierstall, and A. M. Minor, IPAP Conf. Ser. **7**, 380 (2006).

[21] A. Martin, N. Loh, C. Hampton, R. Sierra, F. Wang, A. Barty, A. Aquila, S. Bajt, M. Barthelmess, C. Bostedt, J. Bozek, N. Coppola, *et al.*, *Dark-field femtosecond imaging of airborne soot with an x-ray free electron laser*, (unpublished).

[22] L. W. Whitehead, G. J. Williams, H. M. Quiney, D. J. Vine, R. A. Dilanian, S. Flewett, K. A. Nugent, A. G. Peele, E. Balaur, and I. McNulty, Phys. Rev. Lett. **103**, 243902 (2009).

[23] B. Abbey, L. W. Whitehead, H. M. Quiney, D. J. Vine, G. A. Cadenazzi, C. A. Henderson, K. A. Nugent, E. Balaur, C. T. Putkunz, A. G. Peele, W. J., and McNultyI., Nat Photon **5**, 420 (2011).

[24] H. M. Quiney and K. A. Nugent, Nat. Phys. **7**, 142 (2011).

[25] J. N. Clark, C. T. Putkunz, E. K. Curwood, D. J. Vine, R. Scholten, I. McNulty, K. A. Nugent, and A. G. Peele, Opt. Lett. **36**, 1954 (2011).

[26] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. Bajt, W. H. Benner, R. A. London, *et al.*, Nat. Phys. **2**, 839 (2006).

[27] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, *et al.*, Nature **470**, 73 (2011).

[28] H. N. Chapman, S. P. Hau-Riege, M. J. Bogan, S. Bajt, A. Barty, S. Boutet, S. Marchesini, M. Frank, B. W. Woods, W. H. Benner, R. A. London, U. Rohner, *et al.*, Nature **448**, 676 (2007).

[29] S. P. Hau-Riege, S. Boutet, A. Barty, S. Bajt, M. J. Bogan, M. Frank, J. Andreasson, B. Iwan, M. M. Seibert, J. Hajdu, A. Sakdinawat, J. Schulz, *et al.*, Phys. Rev. Lett. **104**, 064801 (2010).

[30] B. E. Warren, *X-ray diffraction* (Dover, 1990).

[31] V. Elser, *Noise limits in the assembly of diffraction data* (2007), URL `http://www.citebase.org/abstract?id=oai:arXiv.org:0709.3858`.

[32] R. A. Kirian, T. A. White, J. M. Holton, H. N. Chapman, P. Fromme, A. Barty, L. Lomb, A. Aquila, F. R. N. C. Maia, A. V. Martin, R. Fromme, *et al.*, Acta Cryst. A **67**, 131 (2011).

[33] N. D. Loh, M. J. Bogan, V. Elser, A. Barty, S. Boutet, S. Bajt, J. Hajdu, T. Ekeberg, F. R. N. C. Maia, J. Schulz, M. M. Seibert, B. Iwan, *et al.*, Phys. Rev. Lett. **104**, 225501 (2010).

[34] G. Bortel and M. Tegze, Acta Cryst. A **67**, 533 (2011).

[35] B. Moths and A. Ourmazd, Acta Cryst. A **67**, 481 (2011).

[36] R. Fung, V. Shneerson, D. K. Saldin, and A. Ourmazd, Nat. Phys. **5**, 64 (2008).

[37] G. Geloni, V. Kocharyan, and E. Saldin, J. Mod. Opt. **58**, 1391 (2011).

[38] W. Fawley, J. Frisch, Z. Huang, Y. Jiao, H.-D. Nuhn, C. Pellegrini, S. Reiche, and J. Wu, *Toward TW-level, hard X-ray pulses at LCLS*, LCLS TN-11-3, (2011).

[39] S.-K. Son, H. N. Chapman, and R. Santra, Phys. Rev. Lett. **107**, 218102 (2011).

D 11 Anomalous and Resonant X-ray Scattering: Spin, charge and orbital order [1]

J. Strempfer

Deutsches Elektronensynchrotron

22607 Hamburg

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

X-ray diffraction is a well established method for the investigation of the electronic structure of materials. Thomson scattering and more recently magnetic scattering allow to determine the charge distribution and the magnetic structure, respectively. Over the last decades, the importance of resonance effects, which occur at the absorption edges of the atoms present in the studied materials, has increased tremendously. Nowadays almost in every field of x-ray scattering resonant effects are involved. The increased importance is related to the possibility to gain additional information due to the element and polarization sensitivity. The experimental techniques employing resonant scattering can roughly be divided in two groups: (i) the investigation of structural properties and (ii) electronic properties of materials.

The first large field is dealing with the investigation of structural properties with the help of anomalous effects. In the 1980's a first standard anomalous diffraction application was found in MAD (Multi wavelength anomalous diffraction/dispersion). Nowadays, MAD, which is used to overcome the phase problem in the investigation of large molecules [1], or diffraction anomalous fine structure (DAFS) investigation, which allows to obtain information both on local and long range order but also a combination of both are extensively used. Then there are grazing incident DAFS methods being employed [2] to look at nano-wires. In bulk metallic glasses partial structure factors are determined using anomalous scattering effects [3].

The second field deals with the investigation of electronic properties. In this area, resonant x-ray scattering has developed into a powerful technique during the last years. Resonant scattering effects due to charge, magnetic and orbital order are used to get deeper insight into the charge distribution, magnetic ordering phenomena and orbital arrangements within the unit cell. First resonant magnetic scattering experiments were done on holmium single crystals [4]. Due to the large resonance enhancement of the otherwise weak magnetic signal, the method of resonant exchange scattering could be used as a complementary tool for the investigation of magnetic properties, where until then, neutron diffraction was the only method available for microscopic investigations. The possibility to apply the method to orbital order was first shown by Murakami et al. [5]. The results were discussed quite controversially [6, 7, 8]. It is clear today, that in this field the resonance effects can originate from different perturbations to the local electronic environment and only the consideration of polarization and absorption edges can clarify them. Also, a strong interaction between theoretic modeling and experiments is necessary to obtain meaningful results. Especially, it is necessary to use polarization, azimuthal and energy dependences in order to disentangle the different types of order which can occur in the material under investigation. In this field, normally forbidden reflections occur, when the energy is tuned to the absorption edge (charge, orbital order) or the otherwise small magnetic signal is amplified significantly due to resonance enhancement at the absorption edges. This in turn allows also magnetic structure determination by making use of the strong resonance enhancement for magnetic atoms at the absorption edges. With soft x-ray powder diffraction at the Mn L-edges, due to the strong resonant signal, it is possible to investigate orbital and magnetic order from powder samples [9]. These investigation allow to gain additional information on charge, magnetic and orbital degrees of freedom, which is otherwise not accessible. Detailed reviews of the method are available from Vettier [10] and De Bergevin [11] and Dmitrienko [12].

In the following, we will introduce anomalous or resonant scattering and the formalisms used to describe the electronic transitions in order to provide an understanding of the multiple possibilities this method opens for investigation. Then, examples of the different cases will be presented in order to give an impression of the experimental procedures involved.

2 Resonant X-ray Scattering

Resonant X-ray scattering involves electric multipolar transitions from core electrons into an empty valence shell. These virtual transitions into the intermediate states, which occur when the photon energy is tuned close to the absorption edge, produce an increased sensitivity to the local electronic structure, namely the magnetic moment and the anisotropic charge distribution. Since resonant or anomalous scattering occurs at the absorption edges a short introduction into the phenomena occurring at the edges will be given in the following. In general, the terms anomalous scattering and resonant scattering can be used both more or less interchangeably. Nevertheless, the term anomalous scattering is actually used more within the classical resonant scattering community, where structural properties are investigated, whereas the term resonant scattering is used for the investigation of electronic properties, as described in the introduction. The term *anomalous* was coined originally to describe the behavior of the refractive index as function of the wavelength for visible light in transparent material. There, the variation of the refractive index n with increasing wavelength λ is generally negative. This was called *normal* behavior of dispersion. At absorption edges the slope of the variation is positive and dispersion was said to behave *anomalous* [13]. Anomalous dispersion and anomalous scattering are used now in a more general way and express phenomena which are generally related to resonant scattering effects at the absorption edges. In the following, the term resonant scattering will be used instead, since it is a more appropriate term to describe the dispersion effects appearing at the absorption edges.
For x rays interacting with matter, the refractive index can be written as [11]

$$n = 1 - \delta - i\beta \qquad \text{with } \delta \text{ and } \beta \text{ positive.} \tag{1}$$

The real part describes the tiny changes of the refractive index and is generally smaller than 1. The imaginary part describes absorption in the material. Correction terms $f' + if''$ are added to the atomic scattering factor f at the absorption edges in order to take the resonant processes into account. The refractive index decrement δ and the absorption index β can be directly related to the scattering factor corrections by the following relations [14]:

$$\delta = \rho r_0 \frac{\lambda^2}{2\pi}(Z - f'(0)) \tag{2}$$

$$\beta = \rho r_0 \frac{\lambda^2}{2\pi} r f''(0) \tag{3}$$

where ρ is the density of the material, λ the wavelength and r_0 the classical electron radius.
Investigation of crystalline materials with x rays can yield a huge amount of information, if polarization and energy dependence of the absorption and of the scattered signal are taken into account. X-ray absorption measurements are sensitive to the local environment of the respective atom. Here, e.g. XANES or EXAFS use the forward scattered signal and are thus probing the average over the whole crystal. On the other hand, diffraction is probing the long range order by looking at the in phase signal coming from all atoms within the crystal. Taking into account the absorption edges for the diffraction experiments allows to combine the sensitivity of spectroscopy experiments to local electronic configurations with the spacial information and long range order probing capabilities of x-ray diffraction. In this case, as mentioned before, the so-called anomalous dispersion has to be taken into account in the scattering amplitude, which is done by adding an energy dependent correction factor to the scattering factor of the free atom

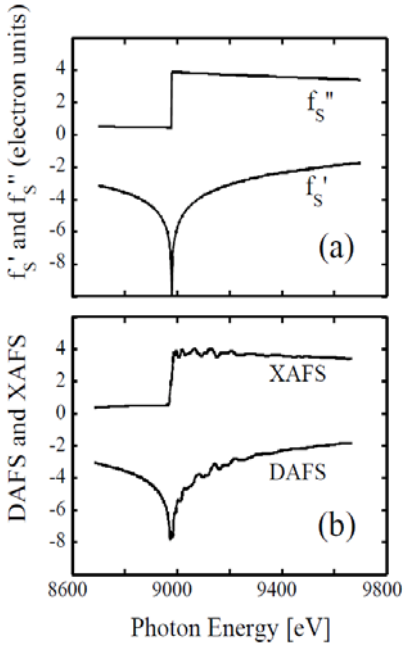


Fig. 1: *(a) Calculated Cromer-Liberman f' and f" for Cu and (b) the respective EXAFS and DAFS spectra. The latter show oscillations over an extended energy range and allow to get information about the atoms and their neighbours [15].*

$f(Q)$. The x-ray scattering amplitude of an atom is written as [16]:

$$f(Q,E) \quad = \quad f_0(Q) + f'(E) + if''(E) + f_{mag}(Q) \tag{4}$$

with the scattering vector $\mathbf{Q} = \mathbf{k}_i - \mathbf{k}_f$, the atomic scattering factor $f_0(Q)$, the energy independent part, which corresponds to Thomson scattering and $f'(E)$ and $f''(E)$ the energy dependent parts, which correspond to the real and imaginary parts of the anomalous scattering or dispersion correction factor (Fig 1). $f_{mag}(Q)$ represents the magnetic scattering factor due to non-resonant magnetic scattering. Normally, $f'(E)$ and $f''(E)$ are polarization independent monotonic functions. Only at the absorption edge, $f''(E)$ exhibits a stepwise increase, whereas $f'(E)$ passes through a deep trough. In case of the forward scattered beam, the dispersion correction factors $f'(E)$ and $f''(E)$ can be transformed into each other by the Kramers-Kronig-relations:

$$f''(\omega) = \frac{2\omega}{\pi}\mathcal{P}\int_0^\infty \frac{f'(\omega')}{(\omega'^2 - \omega^2)}d\omega' \tag{5}$$

$$f'(\omega) = \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\omega' f''(\omega')}{(\omega'^2 - \omega^2)}d\omega' \tag{6}$$

where the principal part of the integral is taken. The coefficients are thus both related, $f''(E)$ directly and $f'(E)$ indirectly, to the absorption cross section by the optical theorem [16, 17]:

$$\sigma_a = 2\frac{4\pi}{k_i}\, Im\left[\boldsymbol{\epsilon}_i^* \cdot f(Q=0)\right] \tag{7}$$

where $\mathbf{k}_i$, $\boldsymbol{\epsilon}_i$ and $\mathbf{k}_f$, $\boldsymbol{\epsilon}_f$ are the wave vector and polarization vector of the incident and scattered beams, respectively. The proof of Eqn. 7 relies on the balance between absorption and flux and is given in Ref. [18].

In 1980, Templeton & Templeton discovered that an anisotropy in the electron distribution around an atom in a compound can lead to intensity at otherwise forbidden positions in reciprocal space when the energy is tuned to the absorption edge [17]. By tuning the incident photon energy close to the absorption edge, core electrons can be excited into the valence states, which are dependent on the chemical bonding and can be distorted by the local environment. This situation cannot be described anymore by scalar expressions as in Eqn. 4 but instead f' and f'' have to be turned into tensors, which are polarization dependent. This so called anisotropy of the tensor of susceptibility (ATS) or Templeton scattering is due to the presence of symmetry elements with translational components. This again gives rise to non-vanishing structure factor tensors violating extinction rules of the crystal. It turned out that this effect can be used to very sensitively probe the local atomic environment. Besides this pure crystalline resonances at normally forbidden reflections, also other effects can lead to resonances which can be due to changes in the electron distribution induced by various correlation effects in the crystal like charge order, orbital order or magnetic order influencing the local symmetry. As we will see, also these effects have been investigated very successfully by resonant scattering in the past years. In dipole approximation, resonant scattering can be described by a symmetric tensor of second rank [13]:

$$f_{res} = f' + f'' = r_0 \sum_{\alpha,\beta} \boldsymbol{\epsilon}'^{*}_{\alpha} \boldsymbol{\epsilon}_{\beta} S^{\beta\alpha} \tag{8}$$

The sum is taken over the Cartesian coordinates α and β, and the polarization of the incident $\boldsymbol{\epsilon}_\beta$ and scattered $\boldsymbol{\epsilon}'_\alpha$ beam are considered in the sum. The detailed form of $S^{\beta\alpha}$ will be developed in the next section.

3 The Resonant Scattering Amplitude

Interaction between photons and electrons in the scattering process can be described by a Hamiltonian which contains the vector potential **A**. There are terms linear in **A** and terms quadratic in **A** [19]. The scattering amplitude is obtained by second order perturbation theory, where the first order term is due to the quadratic terms in **A** and the second order term is arising from the linear terms in **A**. The second term contains energy dependent denominators. The first order term contributes to non-resonant charge and magnetic scattering and will not further be treated here. The second order term, which is responsible for resonant scattering can be written as [20]:

$$\begin{aligned} f_{res} = r_0 \boldsymbol{\epsilon}'^{*}_{\alpha} \boldsymbol{\epsilon}_{\beta} \sum_a p_a \Bigg\{ & \frac{1}{m} \sum_c \left(\frac{E_a - E_c}{\hbar\omega} \right) \frac{< a|\mathbf{O}^{\alpha *}|c >< c|\mathbf{O}^{\beta}|a >}{E_a - E_c + \hbar\omega - i\frac{\Gamma}{2}} \\ & - \frac{1}{m} \sum_c \left(\frac{E_a - E_c}{\hbar\omega} \right) \frac{< a|\mathbf{O}^{\beta}|c >< c|\mathbf{O}^{\alpha *}|a >}{E_a - E_c - \hbar\omega} \Bigg\} \end{aligned} \tag{9}$$

where E_a and E_c are the energy of the initial state $|a>$ and the intermediate state $|c>$, respectively, and Γ is the life time of the intermediate state. $\hbar\omega$ is the photon energy and m the electron mass. The operator O^α is given by the expression:

$$\mathbf{O}^{\alpha}(\mathbf{k}) = \sum_j e^{i\mathbf{k}\cdot\mathbf{r}_j} \left[\mathbf{p}_j^{\alpha} - i\hbar(\mathbf{k} \times \mathbf{s}_j)^{\alpha} \right] \tag{10}$$

which contains the momentum $\mathbf{p}_j$ and the spin $\mathbf{s}_j$ of the electron. The last term in Eqn. 9 can be neglected at resonance compared to the first one, where the denominator is close to

zero. In Eqn. 10, the spin dependent part can be neglected since the transition probabilities in the x-ray regime of the electric multipoles dominate the magnetic transition probabilities. The exponential in Eqn. 10 can be series expanded and the resonant scattering factor decomposed into electric multipole transitions. We finally end up with the following equation for the tensor $S^{\beta\alpha}$ in Eqn. 8:

$$S^{\beta\alpha} = \frac{1}{\hbar\omega}\frac{m}{\hbar^2}\sum_c \left(\frac{(E_a - E_c)^3 \langle a|M^\alpha|c\rangle \langle c|M^\beta|a\rangle}{E_a - E_c + \hbar\omega - i\frac{\Gamma}{2}} \right) \tag{11}$$

where

$$M^{\alpha(\beta)} = r\left(1 - \frac{1}{2}ik^{\alpha(\beta)}\cdot r\right) \tag{12}$$

is the expansion of the photon field up to electric dipole (E1) and quadrupole (E2) terms. Eqn. 11 can be decomposed in a sum of three terms consisting of E1 and E2 contributions[20, 21]:

$$S^{\beta\alpha} = \sum_{\alpha\beta} f^{dd}_{\alpha\beta} - \frac{i}{2}\sum_{\alpha\beta\gamma}\left(k^a_\gamma f^{dq}_{\alpha\beta\gamma} - k^c_\gamma f^{qd}_{\alpha\beta\gamma}\right) + \frac{1}{4}\sum_{\alpha\beta\gamma\delta} k_\gamma k_\delta f^{qq}_{\alpha\beta\gamma\delta} \tag{13}$$

with f^{dd} the E1-E1 scattering, f^{dq} the E1-E2 scattering and f^{qq} the E2-E2 scattering, where these terms are now separated completely from polarization and wave vector terms and leave only the intrinsic anisotropies of the charge distribution. These can be explored by varying and analyzing the x-ray polarization and varying the wave vector.
If we limit ourselves to E1 transitions, the tensor $S^{\beta\alpha}_s$ describing the E1 resonance can be decomposed in three parts: $S^{\beta\alpha}_s = S_{0s}\delta^{\alpha\beta} + S^{\beta\alpha}_{-s} + S^{\beta\alpha}_{+s}$, with $S_{0s}\delta^{\alpha\beta}$ the spherical part, a traceless symmetric $S^{\beta\alpha}_{+s}$ and an anti-symmetric part $S^{\beta\alpha}_{-s}$. The index s indicates the atom in the unit cell. These terms show different polarization dependences. The spherical part corresponds to the classical anomalous scattering represented by the classical dispersion terms f' and f''. If an atom in a uniaxial environment is considered, we obtain:

$$S^{\beta\alpha}_{-s} = 0 \tag{14}$$

$$S^{\beta\alpha}_{+s} = z^\alpha z^\beta - \frac{1}{3}\delta^{\alpha\beta} \tag{15}$$

Here the antisymmetric part is zero and only the symmetric part contributes. For E1 transitions this results in the following expression [20]:

$$f_{res} = -r_0 S_0(\boldsymbol{\epsilon}^*_f\cdot\boldsymbol{\epsilon}_i) - r_0 S_2\left[(\boldsymbol{\epsilon}^*_f\cdot\mathbf{z})(\boldsymbol{\epsilon}_i\cdot\mathbf{z}) - \frac{1}{3}(\boldsymbol{\epsilon}^*_f\cdot\boldsymbol{\epsilon}_i)\right] \tag{16}$$

If a magnetic moment in an otherwise spherically symmetric environment is considered, the antisymmetric part contributes:

$$S^{\beta\alpha}_{-s} = \boldsymbol{\epsilon}^{\alpha\beta\gamma} z^\gamma_m \tag{17}$$

$$S^{\beta\alpha}_{+s} = z^\alpha_m z^\beta_m - \frac{1}{3}z^2_m\delta^{\alpha\beta} \tag{18}$$

where z_m is the pseudo-vector parallel to the magnetization direction. The resonant scattering amplitude for E1 transitions then becomes [10]:

$$f_{res} = -r_0 S_0(\boldsymbol{\epsilon}^*_f\cdot\boldsymbol{\epsilon}_i) - ir_0 S_1(\boldsymbol{\epsilon}^*_f\times\boldsymbol{\epsilon}_i)\cdot\mathbf{z}_m - r_0 S_2\left[(\boldsymbol{\epsilon}^*_f\cdot\mathbf{z}_m)(\boldsymbol{\epsilon}_i\cdot\mathbf{z}_m) - \frac{1}{3}\mathbf{z}^2_m(\boldsymbol{\epsilon}^*_f\cdot\boldsymbol{\epsilon}_i)\right] \tag{19}$$

The first term contributes to Thomson scattering, the second term corresponds to a tensor of rank 1 with odd time reversal symmetry, describing e.g. spin polarization of the orbitals. The third term corresponds to a tensor of rank 2 and depends not on the sign but only on the direction of the magnetic moment. It is similar to the second term in Eqn. 16 and is responsible for the magnetic linear dichroism [11]. Similar but more complex expressions are found for E1-E2 and E2-E2 transitions [22].
With linear polarized x rays, the multipolar order of a system can be explored. In order to get better physical meaning Eqn. 13 can be reformulated, as a scalar product of irreducible tensors instead of using Cartesian coordinates:

$$f_{res} = \sum_{k,q} (-1)^q T_q^k F_{-q}^k(\omega) \tag{20}$$

where T_q^k contains the information about the polarization and wave vectors and F_{-q}^k represents the properties of the system. The rank k represents the order of the multipoles in the field expansion and q the projection in the local axis system of the resonant ion. The description of the multipolar order terms which are probed in the different polarization channels are detailed in Ref. [21]. Here, we will limit ourselves to E1-E1 processes, which are sufficient to describe scattering by charge, magnetic and orbital order.
For a resonant scattering experiment, the processes of photon absorption, virtual photoelectron excitation, and photon re-emission is coherent throughout the crystal. The incident and diffracted x-ray beams produce therefore normal Bragg diffraction. Using the form factor from Eqn. 4, the structure factor is written as:

$$F(Q,E) = \sum e^{i\mathbf{Q}\cdot\mathbf{R}_j} \left(f_0(Q) + f'(E) + if''(E) + f_{mag}(Q)\right) \tag{21}$$

where $\mathbf{R}_j$ denotes the position of the scattering ion j. In the following, we will look in more detail into the description of anisotropic charge scattering as well as magnetic scattering.

3.1 Templeton scattering

In crystallography, selection rules for general and special sets of positions can be different. The special selection rules can be violated by aspherical charge densities or asymmetric thermal motion of the scatterers. In this case, the general selection rules will be still valid. Nevertheless, if the electronic susceptibility is anisotropic, even the general selection rules may be violated. This then leads to the occurrence of resonant reflections which are due to the Anisotropy of the Tensor of Susceptibility and are thus often called ATS-reflections [17].
Conditions that are violated are screw-axis or glide-plane conditions. If the symmetry operation is described by $T(r) = a + R \cdot r$, χ must be invariant under the transformation $\chi(r) = R \cdot \chi(r') \cdot R^{-1}$ with $r' = T^{-1}r$. Therefore, the f^1 tensor of one of the edge atoms has to be transformed according to the site symmetry to find the f tensors of all the other atoms in the unit cell. In the E1-E1 approximation, this would give [12]:

$$F^{\mathbf{Q}} = \sum_j R_j f^1 R^{-1} e^{i\mathbf{Q}\cdot\mathbf{R_j}} \tag{22}$$

where the summation is over all equivalent resonant atoms in the unit cell. For a glide plane forbidden reflection, the glide plane can be considered normal to the x-axis. Then (x, y, z) is

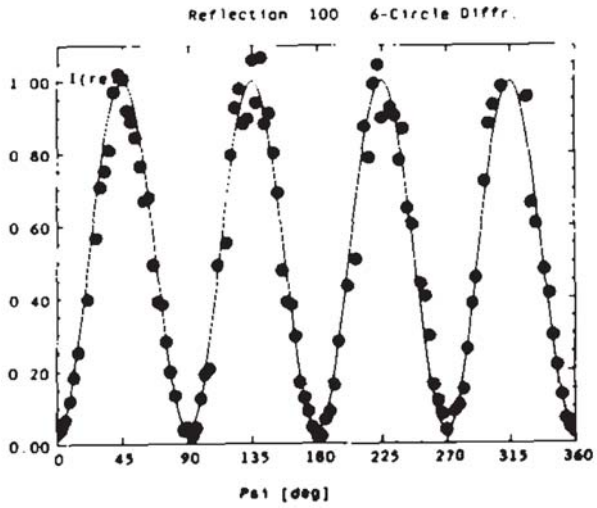


Fig. 2: *Intensity variation as a function of the azimuth angle for the forbidden (3 0 0) reflection of Cu_2O measured at the Cu K absorption edge [24].*

transformed into $(-x, y, z + 1/2)$. The forbidden reflections by the tensor F^{0kl}, if R_m denotes the glide plane transformation:

$$F^{0kl} = e^{-i\pi l} R_m F^{0kl} R_m^{-1} \tag{23}$$

with

$$R_m = R_m^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{24}$$

For forbidden reflections with $l = 2n + 1$, we get $e^{i\pi l} = -1$. In this case, only the traceless part of the tensor remains and it is clear that the Fourier transformation of the density, which is proportional to $\mathrm{Tr}(F^{0kl})$ is zero [23]:

$$F^{0kl} = \begin{pmatrix} 0 & f_1 & f_2 \\ f_1 & 0 & 0 \\ f_2 & 0 & 0 \end{pmatrix} \tag{25}$$

The ATS reflections show a distinct variation with the azimuthal angle, as is the case for electronic resonances as well. Intensities may also change polarization upon scattering from σ to π and vice versa. In Fig. 2, the azimuth dependence of a forbidden (3 0 0) reflection of Cu_2O is shown as an example, revealing a clear 90° periodicity [24].
Some screw axis or glide plane reflections remain forbidden in the E1-E1 approximation. These can be excited if there are E1-E2 or E2-E2 contributions to the ATS tensor [12].

3.2 Magnetic reflections

Magnetic signals could in principle be investigated by non-resonant scattering. Nevertheless, the non-resonant magnetic cross section is 6 to 7 orders of magnitude smaller than the one for Thomson scattering. Additional sensitivity to magnetic properties by x-ray diffraction is given

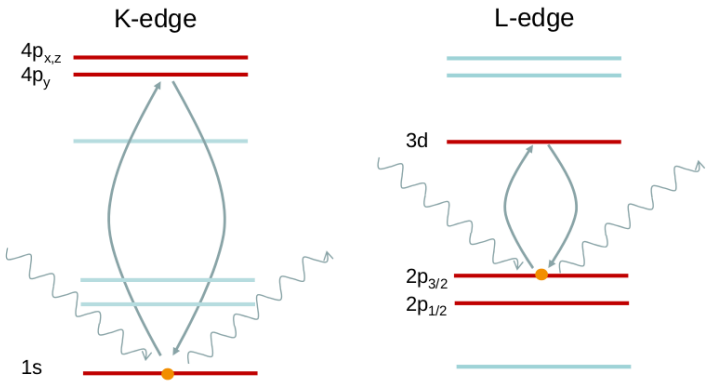


Fig. 3: *Schematics of the electronic states of a Mn^{2+}-atom. Left, the K-edge resonance and right, the L-edge resonances are shown.*

by resonant exchange scattering (XRES) [25]. Core electrons are probing the valence states by virtual transitions into the conduction band. The Pauli principle allows only transition into unoccupied states, which are orbitals with specific magnetic quantum numbers [20]. Exchange interactions between spin moments are polarizing the orbitals, which are then probed by the resonance (see sketch in Fig. 3). This gives a resonant signal due to the magnetic order on top of the non-resonant scattering. The strength of the resonance in the case of XRES strongly depends on the overlap integrals between the core states and the valence states. Typical resonance enhancements are given in Table 1.

Resonant magnetic x-ray scattering was first observed on ferromagnetic Ni at the Ni K-edge in 1985 by Namikawa *et al.* showing a tiny variation of the asymmetric ratio of the intensities with positive and negative fields [26]. At the same time Blume described a possible magnetic resonance [19]. Three years later, the first resonant scattering from antiferromagnetic Ho, performed at the L_3 absorption edge was published by Gibbs *et al.* [4], initiating a rapid growth in this field. The experiment was immediately followed by a thorough theoretical treatment of the resonant magnetic scattering cross section by Hannon *et al.* [25]. The theoretical description was expanded in the following [20, 27]. A detailed description of the effects of polarization with resonant exchange scattering was presented by Hill and McMorrow [22]. The first term and the anti symmetric second term in Eqn. 19 can be written in form of 2×2 matrices for the polarization decomposed in components parallel ($\boldsymbol{\epsilon}_{||}$ or π) and perpendicular ($\boldsymbol{\epsilon}_{\perp}$ or σ) to the scattering plane [22]:

$$\boldsymbol{\epsilon}_f^* \cdot \boldsymbol{\epsilon}_i = \begin{pmatrix} 1 & 0 \\ 0 & \cos 2\theta \end{pmatrix} \tag{26}$$

$$(\boldsymbol{\epsilon}_f^* \times \boldsymbol{\epsilon}_i) \cdot \mathbf{z}_m = \begin{pmatrix} 0 & z_1 \cos\theta + z_3 \sin\theta \\ z_3 \sin\theta - z_1 \cos\theta + & -z_2 \sin 2\theta \end{pmatrix} \tag{27}$$

where z_1, z_2 and z_3 are the components of the magnetic moment along the reference vector of the crystal. A sketch of the scattering configuration is shown in Fig. 4. The first term (Eq. 26), describing Thomson scattering, connects only states for which the polarization is unchanged, whereas the second term (Eq. 27) allows $\sigma - \pi'$ as well as $\pi - \pi'$ scattering, whereas $\sigma - \sigma'$ scattering is forbidden. The third term in Eqn. 19 describes second order terms in $\mathbf{z}_m$ which for example describe second harmonic reflections if incommensurate reflections are investigated.

Table 1: *Resonance enhancement of the most relevant transitions at absorption edges in the soft and hard x-ray regime for XRES ([10, 34]). The resonance enhancement is given in orders of magnitude, where 'weak' corresponds to a factor* < 10*, 'medium' to a factor of about* 10^2 *and 'strong' to a factor* $> 10^3$*.*

Elements	Edge	Transitions	States	Energy range [keV]	Resonance strength
3d	K	E1, E2	1s $\rightarrow$ 4p,3d	5 - 9	weak
3d	L_2, L_3	E1	2p $\rightarrow$ 3d	0.4 - 1.0	strong
4d	L_2, L_3	E1	2p $\rightarrow$ 4d	2.5 - 3.5	strong
4f	L_2, L_3	E1, E2	2p $\rightarrow$ 5d, 4f	6 - 10	medium
4f	M_2, M_3	E1, E2	3p $\rightarrow$ 5d, 4f	1.3 - 2.2	medium
4f	M_4, M_5	E1	3d $\rightarrow$ 4f	0.9 - 1.6	strong
5f	L_2, L_3	E1, E2	3d $\rightarrow$ 6d, 5f	16.3 - 22	medium
5f	M_4, M_5	E1	3d $\rightarrow$ 5f	3.3 - 3.9	strong

It becomes clear from Eqn. 27 that in the $\sigma - \pi'$-channel, magnetic scattering can be clearly distinguished from charge scattering and also moment directions can be determined through polarization analysis.

The magnitude of the resonance effect is varying significantly for different edges as well as for different multipolar transitions. In general, E1 transitions are stronger than E2 transitions. Depending on the valence orbital involved in the transition and the effect probed, i.e. charge order, orbital order or magnetic order, a different strength of the resonance enhancement is observed. In Table 1, some of the most important transitions are shown together with the energy range and the resonance enhancement for XRES. For transition elements, the resonance enhancement at the K-edges in the hard x-ray regime is very small, but large in the soft x-ray region for the 3d L-edges [28] and in the lower hard x-ray region for the 4d L-edges [29]. Transition metals are extremely important for thin films and nanostructures, where effects at the interface between different layers can be investigated by tuning the energy to the absorption edges of the different materials. Unfortunately, the wavelength at these energies is already very large and allows only to reach low indexed superlattice reflections (see Ref. [30] for a review). Resonance enhancements of typically up to 2 orders of magnitude are observed at the rare earth L-edges [31]. In the soft x-ray range scattering at the M_4 and M_5 edges is again important for nanostructured materials [32]. At the M-edges of the actinides, the resonance enhancement in XRES can become up to 7 orders of magnitude [33]. Due to the strong resonance enhancement and modern synchrotron sources it is possible now to perform resonant scattering experiments even on powder samples [9].

3.3 Polarization scans

Polarization analyzers for the scattered beam in the conventional hard x-ray regime (3-15 keV) allows a reliable determination of the polarization of the diffracted signal. The linear polarization of the incident x rays can be varied at the same time very reliably using phase retarder setups with diamond phase plates [35]. This allows to perform a polarization analysis of the diffracted intensity as a function of the incident linear polarization. This so called full polarization analysis, was first successfully applied in the x-ray regime for resonant magnetic scattering

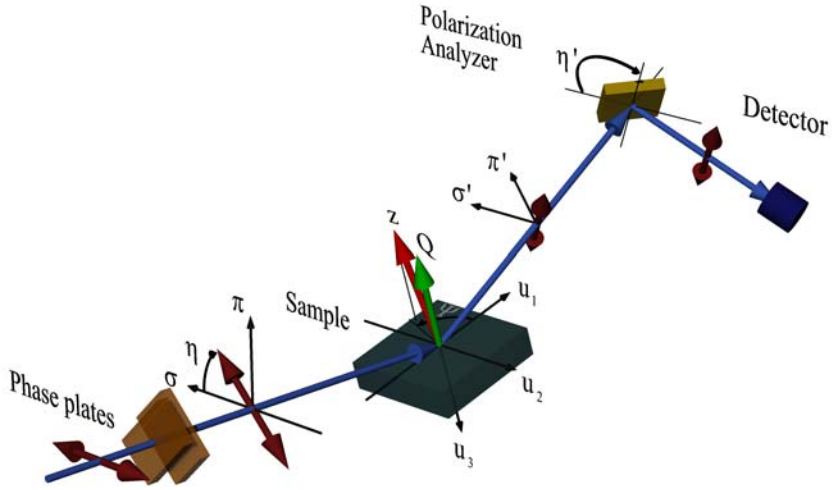


Fig. 4: *Scattering configuration with phase plates and analyzer. The phase plates allow a continuous rotation (η) of the incident polarization as well as generation of left and right circular polarization. For the sample, the scattering vector* **Q** *and the uniaxial vector* **z** *is shown, which rotates around* **Q** *with the azimuth angle* Ψ*. At the sample, the reference vectors* $\mathbf{u}_1$*,* $\mathbf{u}_2$ *and* $\mathbf{u}_3$ *are shown. The scattered signal is analyzed by the analyzer, which can be rotated around the diffracted beam (η'). By scanning the analyzer around the diffracted beam, the diffracted polarization can be determined. Polarization components perpendicular, σ(σ'), and parallel, π(π'), to the diffraction plane are shown for the incident (diffracted) x-ray beam.*

by Mazzoli *et al.* in 2007 [36]. Experimentally, the incident linear polarization is rotated by an angle η and the scattered intensity is analyzed as function of polarization angle η' (see Fig. 4). The incident linear polarization is then varied systematically. A sketch of the experimental setup for full polarization analysis is shown in Fig. 4. The method has become the standard method for the investigation of electronic order and is a basic requirement of any resonant scattering experiment.

In the soft x-ray regime, resonant scattering in the recent years became more and more important for the transition metal L-edges. Here, the generation of variable incident polarization is established already since many years by the use of special insertion devices. Polarization analysis on the contrary is difficult, since integer Bragg reflections cannot be reached at these energies anymore. Nevertheless, graded multilayers were used to perform polarization analysis in the soft x-ray range and full polarization has been performed successfully [37].

In order to describe polarization scans, the Poincaré-Stokes nomenclature is used to describe the polarization states of the x rays. When plane waves with electric field amplitudes E_σ and E_π are used, the Poincaré-Stokes parameters are given by [27]:

$$P_1 = (|E'_\sigma|^2 - |E'_\pi|^2)/P_0 \quad (28)$$

$$P_2 = 2\mathrm{Re}(E'^*_\sigma E'_\pi)/P_0 \quad (29)$$

$$P_3 = 2\mathrm{Im}(E'^*_\sigma E'_\pi)/P_0 \quad (30)$$

with P_1 and P_2 describing the linear polarization states, P_3 the circular polarization state and $P_0 = |E_\sigma|^2 + |E_\pi|^2$. The polarization state of the scattered x rays can be obtained using the expression [36]:

$$I(\eta, \eta') = P_0/2\,(1 + P'_1(\eta)\cos 2\eta' + P'_2(\eta)\sin 2\eta') \quad (31)$$

where η and η' are as shown in Fig. 4. Several rocking curves of the polarizer crystal are collected as a function of η' for a given incident polarization direction η. The integrated intensities are then used to determine P_0, P_1' and P_2'.
The fitting of P_1' and P_2' as function of the incident polarization makes it for example possible to disentangle charge order and orbital order like it was shown for a complex material like magnetite [38], which has attracted some considerable experimental efforts in the recent years.

4 Resonant scattering due to spin, charge and orbital order

Since resonant scattering is sensitive to small changes in the electron distribution around the atom the energy is tuned for, the resonant signal can originate from a variety of influences acting on the atom. The task in an experiment is now to disentangle or determine the origin of the resonant signal. This is achieved by a combination of different experimental approaches: (1) The temperature dependence of the scattered intensity is compared to results obtained through other methods concerning phase transition temperatures or order parameters. (2) Polarization analysis of the diffracted intensity allows a first guess on the nature of the scattered signal according to Table 2. Charge order scatters only in the $\sigma - \sigma'$ and $\pi - \pi'$ channels, whereas for magnetic order scattering in the $\sigma - \sigma'$ channel is forbidden. From orbital or quadrupolar order, scattering in all channels is possible. (3) Investigation of the scattered polarization as function of azimuth and the incident polarization is necessary to determine the symmetry of the scattering amplitude. This allows to fit different models to the results and disentangle possible mixed effects.

Table 2: *Different types of multipolar order which are scattering in the E1-E1 channel, the multipole order k according to Eqn. 20 and the relevant polarization channels. The $k = 1$ and 2 processes are depending on Ψ.*

Order k	Multipole	$\sigma - \sigma'$	$\sigma - \pi'$	$\pi - \sigma'$	$\pi - \pi'$
0	Electric charge	$\neq 0$	$= 0$	$= 0$	$\neq 0$
1	Magnetic dipole	$= 0$	$\neq 0$	$\neq 0$	$\neq 0$
2	Electric quadrupole	$\neq 0$	$\neq 0$	$\neq 0$	$\neq 0$

Some systems are quite complex and therefore show not only one but several types of ordering. This can lead even to interference effects between charge, orbital and magnetic scattering at the same reflection. Thus, interpretation of the data is not straightforward but requires a disentanglement of the contributions through polarization analysis, azimuth dependences and polarizations scans or variation of temperature and photon energy. Such effects have been observed for example in GdB_4 for interference between Templeton scattering and XRES [39] or in Fe_2OBO_3 where isotropic and anisotropic charge order and Thomson scattering are interfering [40]. Interference can occur not only between resonant scattering contributions but also between resonant and non-resonant contributions, where also non-resonant scattering may show a variation of intensity with polarization or azimuth, as is the case for resonant magnetic diffraction.

4.1 Charge and orbital order

Charge ordering, i.e. periodically slightly varying charges at equivalent atom positions in the unit cell, can be probed by resonant scattering. Due to the slightly different charges at atom

positions with almost identical site symmetries, tiny shifts in energy can result in differences of the resonance energies. This again can cause the atomic structure factors of the different atoms to almost subtract (Eqn. 21), which causes a significant amplification of the small effect [41]. It is often very difficult to separate charge order from other sources of resonant scattering like orbital order associated with Jahn-Teller (JT) distortions [42]. It is therefore very important to describe the expected results by theoretical models and compare the simulations with the experimental data.

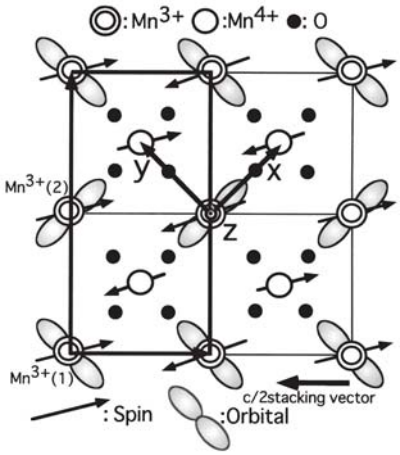


Fig. 5: *Schematic view of the charge, spin and orbital ordering in $La_{0.5}Sr_{1.5}MnO_4$ [5]*

Orbital order was first investigated using resonant scattering by Murakami *et al.* on the layered perovskite $La_{0.5}Sr_{1.5}MnO_4$ [5] and on the three-dimensional perovskite $LaMnO_3$ [43], using resonant scattering at the Mn K-edge. It was proposed that the anisotropy of the Mn^{3+} e_g wave function in the orbital ordered state (Fig. 5) leads to Templeton scattering, since the form factors of Mn^{3+} and Mn^{4+} are different. The scattered intensity of the orbital reflection (3/4 3/4 0) is observed in the $\sigma - \pi'$-channel and shows a variation in azimuth described by:

$$I(\theta, \psi) \propto I_{res}^{OO} \cos^2 \theta_B \cos^2 \Psi \tag{32}$$

This investigation triggered a significant exploration of the method in this direction and opened a new field for the application of resonant x-ray scattering. It is described by the second term in Eqn. 16. RXS at the K-edge involves virtual transitions into the unoccupied 4p orbitals (Fig 3) and sensitivity to 3d orbital order is indirect. Therefore, the K-edges resonance is very sensitive to 4p-band structure effects mediated by 3d-4p Coulomb interactions, which directly influence the electron distribution at the 4p levels. That means, orbital order is measured through the associated cooperative Jahn-Teller (JT) distortions [44]. It was predicted theoretically that L-edge diffraction would provide information on the type of orbital ordering as well as permitting "the effects of orbital ordering and Jahn-Teller ordering to be detected and distinguished from one another" [6]. By L-edge resonances, accessible at soft x-ray energies, 3d-levels, which means the ordered orbitals, can be probed directly. Experiments in the soft x-ray range were only conducted a few years later, but it could be shown that indeed a measurement of the orbital order is possible but suffers still from interaction with JT distortions as well as with magnetic order [7, 45]. It was pointed out by Di Matteo, that a direct probing of orbital order may not be possible, but that it always will be measured indirectly through the JT-effect [8].

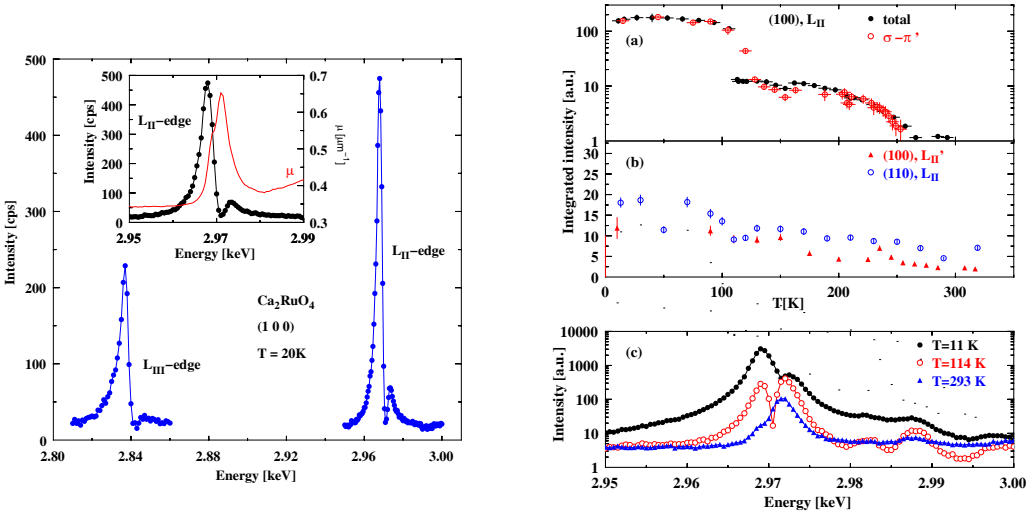


Fig. 6: *Left: resonances of the magnetic (1 0 0) reflection of Ca_2RuO_4 at the L_2 and L_3 absorption edges. The inset shows the L_2 resonance with respect to the absorption edge. Right: The temperature dependence of the intensity of the (1 0 0) resonant reflection at the L_3-edge is shown for (a) the left feature in the energy scan shown in (c) and (b) of the right feature of the energy scan [29].*

4.2 Magnetic order

Resonant magnetic diffraction is mostly done at L-edges or M-edges, where the largest resonance enhancement is obtained (see Table 1). The square root of the resonant intensity is proportional to the order parameter, $M \propto \sqrt{I}$ and critical parameters can be determined very accurately [31]. In compounds with different magnetic elements, the contributions can be separated by tuning to different absorption edges, as has been done for $Gd_xEu_{1-x}S$ [46].

As an example of a system, where magnetic order but also other types of ordering are found, we will look at the 4d transition metal oxide Mott insulator Ca_2RuO_4. Several controversial predictions have been made for the ordering of the 4d t_{2g} Ru orbitals in this system, which provide a strong motivation for the investigation by resonant scattering. A strong resonance enhancement is found at the Ru L_2 (2.9685 keV) and L_3 (2.837 keV) absorption edges at the (1 0 0) reflection in the $\sigma - \pi'$ channel in the antiferromagnetically ordered phase. This is attributed to magnetic order, which has been reported before by neutron powder diffraction. The resonances below the magnetic transition temperature of T_N=110 K are shown in Fig 6 (left). In Fig 6 (right,a), the temperature dependence of the magnetic reflection intensity is shown in logarithmic scale. It can be seen that intensity is remaining above T_N, which only vanishes at T_{OO}=260 K. This intensity is attributed to orbital order. The feature at the right side of the resonance in Fig 6 (right,c) does not vanish at T_{OO} but remains visible up to the metal-insulator transition T_{MI}=357 K. It shows the same temperature dependence as the resonant scattering observed at the (1 1 0)-reflection, which is attributed to tilt order of the oxygen octahedra.

Through the form of the magnetic cross section (Eqn. 27), resonant magnetic scattering allows to determine magnetic moment directions very reliably. In the superconducting compound $RuSr_2GdCu_2O_8$, which shows superconductivity and magnetism simultaneously in the same unit cell, it was possible to clarify ambiguous results from neutron diffraction by an azimuthal scan of the resonant intensity of the magnetic (1/2 1/2 1/2) superlattice reflection. From Fig. 7

it is possible to deduce a magnetic moment direction lying close to the (1 0 2) direction, instead of (0 0 1) as originally assumed, by fitting the expression [47]:

$$I^{\sigma-\pi'}_{\frac{1}{2}\frac{1}{2}\frac{1}{2}} = \left|\sin\alpha\cos\theta\cos(\psi-\psi_0)+\cos\alpha\sin\theta\right|^2 \tag{33}$$

where α is the angle between the magnetic moment and the scattering vector.

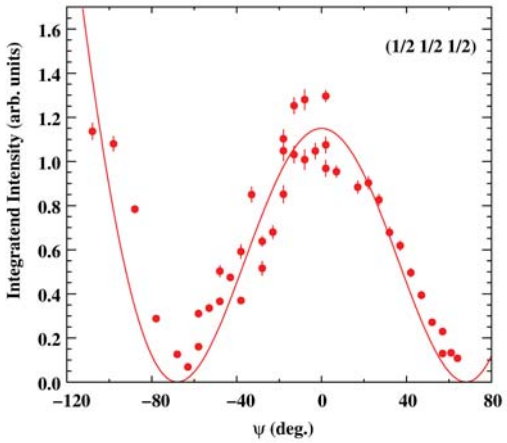


Fig. 7: *Azimuth dependence of the resonant magnetic scattering intensity for the (1/2 1/2 1/2) reflection of $RuSr_2GdCu_2O_8$ [47]. The solid points are measured data, whereas the line is a fit of the theory for XRES.*

References

[1] V. Favre-Nicolin, S. Bos, J. E. Lorenzo, P. Bordet, W. Shepard, and J. L. Hodeau, J. Appl. Cryst. **33**, 52 (2000).

[2] S. Grenier, M. G. Proietti, H. Renevier, L. Gonzlez, J. M. Garca, and J. Garca, Europhys. Lett. **57**, 499 (2002).

[3] S. Hosokawa and J.-F. Bérar, AIP Conf. Proc. **879**, 1743 (2007).

[4] D. Gibbs, D. Harshman, E. Isaacs, D. McWhan, D. Mills, and C. Vettier, Phys. Rev. Lett. **61**, 1241 (1988).

[5] Y. Murakami, H. Kawada, H. Kawata, M. Tanaka, T. Arima, Y. Moritomo, and Y. Tokura, Phys. Rev. Letters **80**, 1932 (1998).

[6] C. W. M. Castleton and M. Altarelli, Phys. Rev. B **62**, 1033 (2000).

[7] S. S. Dhesi, A. Mirone, C. D. Nadai, P. Ohresser, P. Bencok, N. B. Brookes, P. Reutler, A. Revcolevschi, A. Tagliaferri, O. Toulemonde, and G. van der Laan, Phys. Rev. Letters **92**, 056403 (2004).

[8] S. di Matteo, J Physics: Conf. Series **190**, 012008 (2009).

[9] U. Staub, M. Garcia-Fernandez, Y. Bodenthin, V. Scagnoli, R. A. de Souza, M. Garganourakis, E. Pomjakushina, and K. Conder, Phys. Rev. B **79**, 224419 (2009).

[10] C. Vettier, J. Elec. Spectr. Rel. Phenomena **117-118**, 113 (2001).

[11] F. de Bergevin, in *X-ray and Neutron Reflectivity: Principles and Applications*, edited by J. Daillant and A. Gibaud (Springer, 2009), pp. 1–57.

[12] V. E. Dmitrienko, K. Ishida, A. Kirfel, and E. N. Ovchinnikova, Acta Cryst. A **61**, 481 (2005).

[13] D. H. Templeton, in *Resonant Anomalous X-ray Scattering Theory and Applications*, edited by G. Materlik, C. J. Sparks, and K. Fischer (Elsevier Science B.V., 1994).

[14] A. Thompson, D. Attwood, and E. G. et al., *X-ray data booklet*, http://xdb.lbl.gov/.

[15] L. B. Sorensen, J. O. Cross, M. Mewville, B. Ravel, J. J. Rehr, H. Stragier, C. E. Bouldin, and J. C. Woicik, in *Resonant anomalous x–ray scattering*, edited by C. S. G. Materlik and K.Fischer (Elsevier Science B.V., 1994), pp. 231–256.

[16] R. W. James, *The Optical Principles of the Diffraction of X-rays* (G. Bell And Sons Limited, 1962), chap. 4.

[17] D. H. Templeton and L. K. Templeton, Acta Cryst. A **36**, 237 (1980).

[18] J. D. Jackson, *Klassische Elektrodynamik* (de Gruyter, 1983).

[19] M. Blume, J. Appl. Phys. **57**, 3615 (1985).

[20] M. Blume, in *Resonant anomalous x–ray scattering*, edited by C. S. G. Materlik and K.Fischer (Elsevier Science B.V., 1994), pp. 231–256.

[21] S. di Matteo, Journal of Physics: Conference Series **211**, 012008 (2010).

[22] J. P. Hill and D. F. McMorrow, Acta Cryst. A **52**, 236 (1996).

[23] V. E. Dmitrienko, Acta Cryst. A **39**, 29 (1983).

[24] A. Kirfel, A. Pecov, and K. Eichhorn, Acta Cryst. A **47**, 180 (1991).

[25] J. Hannon, G. Trammel, M. Blume, and D. Gibbs, Phys. Rev. Lett. **61**, 1245 (1988).

[26] K. Namikawa, M. Ando, T. Nakajima, and H. Kawata, J. Phys. Soc. Japan **54**, 4099 (1985).

[27] M. Blume and D. Gibbs, Phys. Rev. B **37**, 1779 (1988).

[28] S. B.Wilkins, P. D. Hatton, M. D. Roper, D. Prabhakaran, and A. T. Boothroyd, Phys. Rev. Letters **90**, 187201 (2003).

[29] I. Zegkinoglou, J. Strempfer, C. Nelson, J. Hill, J. Chakhalian, C. Bernhard, J. Lang, G. Srajer, H. Fukazawa, S. Nakatsuji, Y. Maeno, and B. Keimer, Phys. Rev. Lett. **95**, 136401 (2005).

[30] G. van der Laan, Current Opinion in Solid State and Materials Science **10**, 120 (2006).

[31] T. Brückel, D. Hupfeld, J. Strempfer, W. A. Caliebe, K. Mattenberger, A. Stunault, N. Bernhoeft, and G. I. McIntyre, Eur. Phys. J. B **19**, 475 (2001).

[32] E. Weschke, H. Ott, E. Schierle, C. Schüssler-Langeheine, D. Vyalikh, G. Kaindl, V. Leiner, M. Ay, T. Schmitte, H. Zabel, and P. J. Jensen, Phys. Rev. Letters **93**, 157204 (2004).

[33] E. D. Isaacs, D. B. McWhan, C. Peters, G. E. Ice, D. P. Siddons, J. B. Hastings, C. Vettier, and O. Vogt, Phys. Rev. Letters **62**, 1671 (1989).

[34] T. Brückel, Lecture Notes of the 36^{rd} IFF Spring School "Magnetism goes Nano" p. B5.1 (2005).

[35] C. Detlefs, M. S. del Rio, and C. Mazzoli, arXiv **1106.4446** (2011).

[36] C. Mazzoli, S. B. Wilkins, S. D. Matteo, B. Detlefs, C. Detlefs, V. Scagnoli, L. Paolasini, and P. Ghigna, Phys. Rev. B **76**, 195118 (2007).

[37] U. Staub, V. Scagnoli, Y. Bodenthin, M. Garcia-Fernandez, R. Wetter, A. M. Mulders, H. Grimmere, and M. Horisberger, J. Synchrotron Rad. **15**, 469 (2008).

[38] S. R. Bland, B. Detlefs, S. B. Wilkins, T. A. W. Beale, C. Mazzoli, Y. Joly, P. D. Hatton, J. E. Lorenzo, and V. A. M. Brabers, J. Phys.: Condens. Matter **21**, 485601 (2009).

[39] S. Ji, C. Song, J. Koo, K.-B. Lee, Y. J. Park, J. Y. Kim, J.-H. Park, H. J. Shin, J. S. Rhyee, B. H. Oh, and B. K. Cho, Phys. Rev. Letters **91**, 257205 (2003).

[40] S. R. Bland, M. Angst, S. Adiga, V. Scagnoli, R. D. Johnson, J. Herrero-Martn, and P. D. Hatton, Phys. Rev. B **82**, 115110 (2010).

[41] Y. Joly, J. E. Lorenzo, E. Nazarenko, J.-L. Hodeau, D. Mannix, and C. Marin, Phys. Rev. B **78**, 134110 (2008).

[42] V. Scagnoli, U. Staub, M. Janousch, A. M. Mulders, M. Shi, G. I. Meijer, S. Rosenkranz, S. B. Wilkins, L. Paolasini, J. Karpinski, S. M. Kazakov, and S. W. Lovesey, Phys. Rev. B **72**, 155111 (2005).

[43] Y. Murakami, J. P. Hill, D. Gibbs, M. Blume, I. Koyama, M. Tanaka, H. Kawata, T. Arima, Y. Tokura, K. Hirota, and Y. Endoh, Phys. Rev. Letters **81**, 582 (1998).

[44] M. v. Zimmermann, C. S. Nelson, J. P. Hill, D. Gibbs, M. Blume, D. Casa, B. Keimer, Y. Murakami, C.-C. Kao, C. Venkataraman, T. Gog, Y. Tomioka, *et al.*, Phys. Rev. B **64**, 195133 (2001).

[45] S. B. Wilkins, N. Stoji, T. A. W. Beale, N. Binggeli, C. W. M. Castleton, P. Bencok, D. Prabhakaran, A. T. Boothroyd, P. D. Hatton, and M. Altarelli, Phys. Rev. B **71**, 245102 (2005).

[46] D. Hupfeld, W. Schweika, J. Strempfer, W. A. Caliebe, U. Köbler, K. Mattenberger, G. J. McIntyre, F. Yakhou, and T. Brückel, Eur. Phys. J. B **26**, 273 (2002).

[47] B. Bohnenbuck, I. Zegkinoglou, J. Strempfer, C. Nelson, H.-H. Wu, C. Schüssler-Langeheine, M. Reehuis, E. Schierle, P. Leininger, T. Herrmannsdörfer, J. Lang, G. Srajer, *et al.*, Phys. Rev. Letters **102**, 037205 (2009).

E 1 Superconductivity [1]

Y. Xiao and Y. Su
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Superconductivity is an exotic state of matter that has fascinated generations of scientists ever since its discovery in mercury in 1911 [1]. After more than 100 years, there are whole classes of superconducting materials that we still do not fully understand. In particular, understanding the mechanism of high temperature superconductivity in cuprates, discovered in 1986 by Bednorz and Müller [2], and in recently discovered iron-based superconductors [3] has remained as one of the hardest tasks in condensed matter physics.

In this lecture note, the fundamentals of superconductivity, as commonly covered in most textbooks on solid state physics, will be given first. Detailed descriptions on the phenomenon of superconductivity, phenomenological theories i.e. the London theory and Ginzburg-Landau theory and the landmark microscopic BCS theory can be found in this section. A brief overview of large classes of the known superconducting materials, including both conventional and unconventional superconductors will be given as well. The high temperature superconductivity is a vast and always rapidly moving field, it is neither in our intention nor possible to present a comprehensive review on the current status as well as a thorough description of the underlying physics. Therefore, only some important aspects and basic understanding of the recently discovered iron-based superconductors will be discussed here. The choice of the covered topic is more or less based on our own research activities in this field with the main focus on the neutron scattering studies.

2 Fundamentals of Superconductivity

2.1 Zero resistivity

In order to explain the electronic property of metals, the Drude theory was developed by Drude in 1900. In the framework of the Drude theory, electrons are treated as classical particles. When electrons are moving through the solid, they will collide with scattering center and change their direction and velocity. In a metal, the conductivity can be defined by the constitutive equation as the proportionality between electrical current density **J** and electric field **E**: **J**=σ**E**. The electrical conductivity σ is given by Drude theory as:

$$\sigma = \frac{ne^2\tau}{m} \tag{1}$$

where τ is the mean life time and m is the effective mass of the conduction electrons. The resistivity ρ is the reciprocal of the conductivity. *i.e.* $\rho = 1/\sigma$ and $\rho \propto \tau^{-1}$. The resistivity is the sum of the contributions from different scattering processes. The scattering mechanisms can be impurity scattering, electron-electron scattering and electron-phonon scattering. Furthermore, these scattering processes act independently and they have different mean life times. The total resistivity reads:

$$\rho = \frac{m}{ne^2}(\tau_i^{-1} + \tau_{e-e}^{-1} + \tau_{e-p}^{-1}) \tag{2}$$

The temperature dependencies of these mean life times are also different. The impurity scattering life time is independent of temperature, while the lifetimes of electron-electron and electron-phonon scattering will exhibit temperature variation. At low temperatures, electron-electron

scattering and electron-phonon scattering are negligible. Therefore, one expects a constant value for resistivity at zero temperature as the residual resistivity.

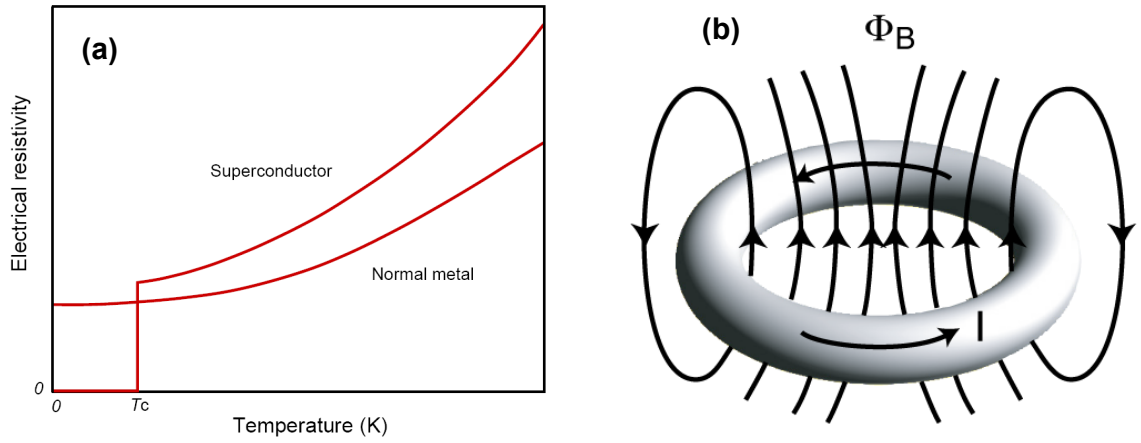


Fig. 1: *(a): Temperature dependence of the electrical resistivity for superconductors and normal metals. (b): The superconducting current maintains through the superconducting ring [4, 5].*

In 1911, Heike Kamerlingh Onnes performed an experiment to test the validity of the Drude theory by measuring the resistivity of mercury at low temperature. Surprisingly, he found the electrical resistance dropped sharply to zero below 4.2 K [1]. Thus the superconductivity was discovered and it represents a new state of matter [4, 5].
As shown in Fig. 1(a), the resistivity of superconductor drops to zero when temperature is below the critical temperature T_c. However, due to the experimental difficulty, we are not be able to measure the zero resistivity. The existence of persistent current in a closed loop of superconducting wire can be considered as the strong evidence of superconductivity. As shown in Fig. 1(b), a circulating current I can be introduced in the superconducting ring. If the superconductor has the zero resistivity, the energy stored in the ring will keep constant and the current will continue flowing in the ring. Experimentally, it was found that almost no detectable decay of the current in superconducting ring for years.

2.2 Meissner effect

Suppose we take a superconductor and place it under magnetic field, the external magnetic field will penetrate into the superconductor if the superconductor is in its normal state, *i.e.* $T > T_c$. Then we will get almost same value of magnetic field inside and outside of superconductor as indicated in Fig. 2(a). When we cool the superconducting sample below T_c in the presence of the same field, the magnetic field will be expelled from the sample. This phenomenon was discovered by Meissner in 1933 and named as Meissner effect [6].
As we known, the superconductor exhibits zero resistivity. By $\mathbf{E} = \rho\mathbf{J}$, we will have $\mathbf{E} = 0$ inside of superconductor. By using the Maxwell equation $\nabla\times\mathbf{E} = -\partial\mathbf{B}/\partial t$, we got $\partial\mathbf{B}/\partial t = 0$. It is also known that magnetic flux density $\mathbf{B}$ is related with magnetic field $\mathbf{H}$ and magnetization of sample $\mathbf{M}$ by $\mathbf{B} = \mu_0(\mathbf{H}+\mathbf{M})$. Eventually, we find that $\partial\mathbf{M}/\partial\mathbf{H} = \chi = -1$ for superconductor. Susceptibility $\chi = -1$ indicated that the superconductors possess perfect diamagnetism. To study the magnetic susceptibility as a function of temperature, we will be able to characterize the

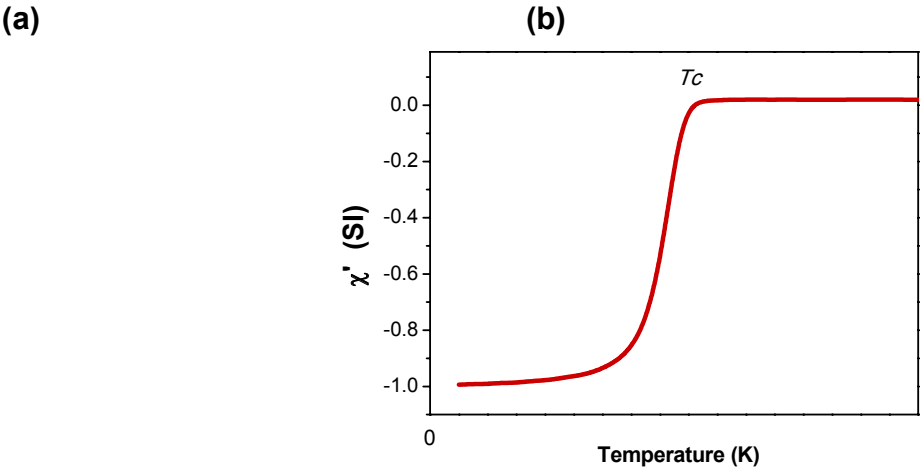


Fig. 2: *(a): The Meissner effect in superconductors. (b): Temperature dependence of susceptibility (χ) of bulk superconductor [4, 5].*

superconducting sample. As shown in Fig. 2(b), we will obtain χ = -1 for bulk superconductor below T_c, which is the solid evidence for Meissner effect.

2.3 London theory

In 1935, London brothers developed the first theory to explain the magnetic properties of superconductors [7]. By applying the two-fluid model, the 1^{st} London equation can be obtained, which relate the superconducting current density **J** with the electric field **E**:

$$\mathbf{E} = \mu_0 \lambda_L^2 \frac{\partial \mathbf{J}}{\partial t} \tag{3}$$

Combination of Eq. (3) with Maxwell equation $\nabla \times \mathbf{E}$ = -$\partial \mathbf{B}/\partial t$, London equation can also be rewritten in terms of magnetic field **B** and superconducting current density **J**, which is called 2^{nd} London equation:

$$\mathbf{B} = -\mu_0 \lambda_L^2 \nabla \times \mathbf{J} \tag{4}$$

In both Eq. (3) and Eq. (4), λ_L is the London penetration depth with the dimension of length,

$$\lambda_L = (\frac{m_e}{\mu_0 n_s e^2})^{1/2} \tag{5}$$

The London equations provide a simple phenomenological model to explain the Meissner effect. It also implies that the magnetic field will only penetrate the surface layer of depth λ_L and the field equals to zero inside the bulk superconductor.

2.4 Ginzburg-Landau theory

In 1930, Landau had developed a theory for second-order phase transition. Many second-order phase transitions can be characterized by an appropriate order parameter, the order parameter shows different value at high temperature disordered state and low temperature order state. For

example, the magnetic order parameter is always used to describe the magnetic phase transition from ferromagnetism to paramagnetism. In 1950, Ginzburg and Landau proposed to describe the superconducting state with a more complex order parameter Ψ, here Ψ is spatially varied, and $|\Psi|^2$ is proportional to the density of super electrons, *i.e.* $|\Psi|^2 = n_s(r)$. Ψ is nonzero in the superconducting state, while it equals zero in normal state above critical temperature T_c [8]. Since superconducting state is a thermal equilibrium state, its thermal dynamic property can be described with free energy density f_s. For temperature close to critical temperature, free energy can be expanded as a function of order parameter $|\Psi|$,

$$f_s(T) = f_n(T) + a(T)|\Psi|^2 + \frac{b(T)}{2}|\Psi|^4 + \cdots \tag{6}$$

where $f_s(T)$ is the free energy density of the normal state, *a* and *b* are the temperature dependent parameters, In order to get minimum for f_s, *b(T)* has to be positive, while *a(T)* can be either positive or negative, corresponding to $T > T_c$ or $T < T_c$, respectively. If we plot the difference of free energy density as a function of Ψ, we will get two different curves for *a(T)* > 0 and *a(T)* < 0. These two curves have different minimum: at Ψ = 0 for $T > T_c$ and at $|\Psi|^2$ = *-a(T)/b(T)* for $T < T_c$.

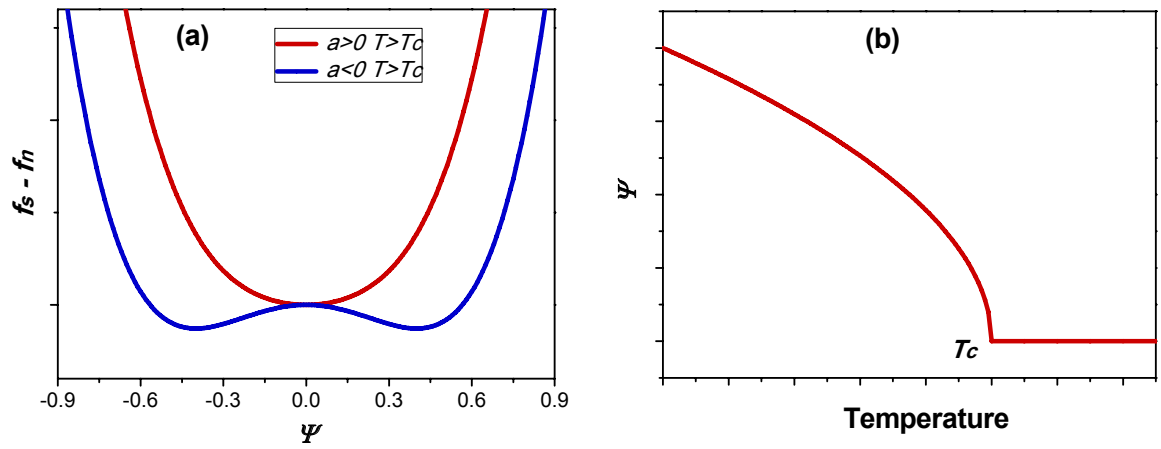


Fig. 3: *(a): Difference between the free energy in the normal and superconducting state as a function of order parameter Ψ. (b): Temperature dependence of order parameter Ψ in superconductor.*

In a spatially inhomogeneous superconductor, the order parameter depends on position. A new term depending on the gradient of $\Psi(r)$ should be included in the free energy. If we also consider the effect of magnetic field $\mathbf{B} = \mu_0\mathbf{H}$, another additional term should also be included. Therefore the free energy of superconductor in the magnetic field is:

$$f_s(T) = f_n(T) + a|\Psi|^2 + \frac{b}{2}|\Psi|^4 + \frac{1}{2m_s}|(-i\hbar\nabla - 2eA)\Psi|^2 + \mu_0\frac{|H|^2}{2} \tag{7}$$

By minimizing the free energy of the system, we can get two Ginzburg-Landau equations:

$$a\Psi + b|\Psi|^2\Psi + \frac{1}{2m_s}(-i\hbar\nabla - 2eA)^2\Psi = 0 \tag{8}$$

$$\mathbf{J}_s = -\frac{2e\hbar i}{i2m_s}(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) - \frac{(2e)^2}{m_s}|\Psi|^2 A \tag{9}$$

Suppose there exists an interface between normal state and superconducting state. By assuming that $\Psi(r)$ is continuous and the boundary condition at $\Psi(0) = 0$, we can solve the first Ginzburg-Landau equation and get $\Psi(r)$:

$$\Psi(r) = \Psi(0)tanh(\frac{r}{\sqrt{2}\xi(T)}) \tag{10}$$

where $\Psi(0)$ is the order parameter far from the interface in the superconducting and $\xi(T)$ is called Ginzburg-Landau coherence length:

$$\xi(T) = \sqrt{\frac{\hbar^2}{2m_s|a(T)|}} \tag{11}$$

Beside of the London penetration depth λ_L, Ginzburg-Landau coherence length $\xi(T)$ is another fundamental length scale associated with superconductivity. The ratio between these two length scales is denoted as Ginzburg-Landau parameter, which is independent of temperature,

$$\kappa = \frac{\lambda(T)}{\xi(T)} \tag{12}$$

Usually, the ratio $\kappa = 1/\sqrt{2}$ is adopted as the criterion to define the type-I and type-II superconductors:

$$\kappa \leq 1/\sqrt{2} \qquad (Type - I) \tag{13}$$

$$\kappa \geq 1/\sqrt{2} \qquad (Type - II) \tag{14}$$

For type-I superconductor, the field inside is zero due to the Meissner effect, when external field is larger than critical field H_c, the superconductivity is destroyed suddenly. While, there are two different critical fields in type-II superconductor: the lower critical field H_{C1} and upper critical field H_{C2}. If external field is smaller than H_{C1}, the sample is perfect diamagnet. If external field exceeds H_{C1} but below H_{C2}, the superconductor enters the so called Shubnikov phase, in which the magnetic flux penetrates the superconductor in the form of vortices [9]. If external field increases further, the vortex cores are getting closer and almost overlap when external field reached upper critical field H_{C2}. The superconductivity will be totally destroyed once the field exceeds H_{C2}.

2.5 Electron-pairing and the BCS theory

As a phenomenological theory, the Ginzburg-Landau theory was quite successful in explaining many physical properties of superconductor. However, it can not explain the microscopic origins of superconductivity. For instance, the physical meaning of the Ginzburg-Landau order parameter was still no clear. In 1957, Bardeen, Cooper, and Schrieffer (BCS) proposed a microscopic theory which can provide the physical interpretation of the nature of order parameter and describe the macroscopic wavefunction of conduction electrons [10]. The key idea of BCS theory is that the crystal lattice phonons can act as the exchange bosons and give an attractive

interaction between the conduction electrons near the Fermi surface. Thus, the pair bound state of electrons is formed and the paired conduction electrons are called "Cooper pair".

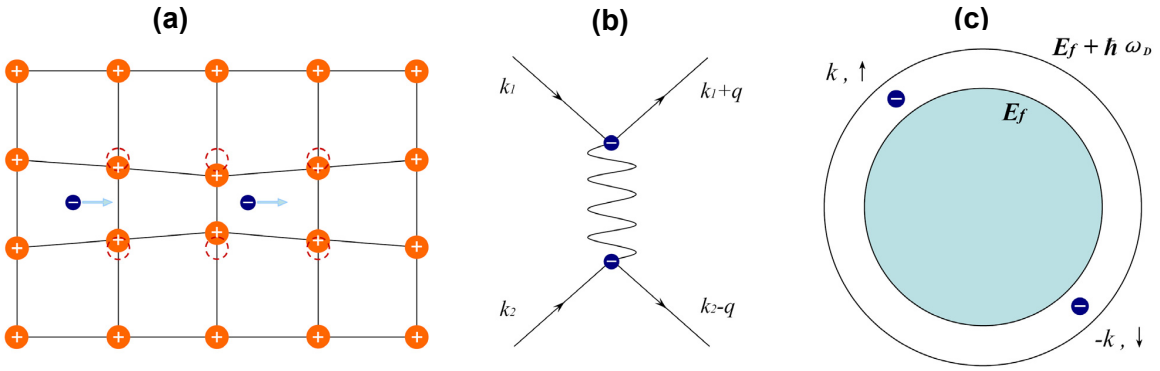


Fig. 4: *(a) Two electrons paired when moving through crystal lattice due to the electron phonon coupling. (b) Interaction of electrons via exchange of boson (crystal lattice phonon). (c) Attractive interaction between two electrons close to the Fermi surface.*

It is well known that bare electrons will repel each other due to the strong electrostatic Coulomb repulsion. However, if we consider the electrons in a medium, say, surrounded by charged ions in a crystal lattice [Fig. 4(a)], the Coulomb interaction will be largely reduced by the screening. Furthermore, the attractive electron-electron interaction is generated via the exchange of virtual exchange bosons such as phonons. As presented in Fig. 4(b), the Feynman diagram illustrated the interaction of electrons via exchange phonons. Because the total wave vector is conserved, we have k_1+k_2 = k_1+q+k_2-q = k_1'+k_2'. Considering a situation that only two additional electrons located outside of the spherical Fermi surface at $T = 0$. The interaction between two additional electrons will take place within the range $\hbar\omega_D$ of Fermi surface, *i.e.* $E_f < E_k < E_f$+$\hbar\omega_D$. To ensure the momentum conservation and to minimize the energy, two additional electrons will pair up as Cooper pair with no center of mass motion, as demonstrated in Fig. 4(c).
The coupling of the electron spins of the Cooper pairs will result in two different total spin, S = 0 or S = 1. Thus the spin wave function can be:

$$\phi = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \qquad (S = 0, singlet) \tag{15}$$

$$or \quad \phi = \begin{cases} |\uparrow\uparrow\rangle \\ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |\downarrow\downarrow\rangle \end{cases} \qquad (S = 1, triplet) \tag{16}$$

Furthermore, based on the distribution of pairing amplitude in k space, the pairing symmetry can be classified as *s, p, d, f*...waves.
By using the language of second quantization, the pairing hamiltonian of singlet superconductor can be given as:

$$\mathcal{H} = \sum_{k,\sigma} \epsilon(k) n_{k\sigma} + \sum_{k,k'} V_{k,k'} c^{\dagger}_{k\uparrow} c^{\dagger}_{-k\downarrow} c_{-k'\downarrow} c_{k'\uparrow} \tag{17}$$

By defining $b_k = \langle c_{-k\downarrow} c_{k\uparrow} \rangle$ and $\triangle_k = -\sum_{k'} V_{k,k'} \langle c_{-k\downarrow} c_{k\uparrow} \rangle$, the model Hamiltonian is expressed as:

$$\mathcal{H} = \mu N + \sum_{k,\sigma} \xi_k c^\dagger_{k\sigma} c_{k\sigma} - \sum_k (\triangle_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} + \triangle^*_k c^\dagger_{-k\downarrow} c^\dagger_{k\uparrow}) \tag{18}$$

The above Hamiltonian can be diagonalized by Bogoliubov-Valatin transformation:

$$c_{k\uparrow} = u^*_k \gamma_{k\uparrow} + v_k \gamma^\dagger_{-k\downarrow} \tag{19}$$

$$c^\dagger_{k\uparrow} = -v^*_k \gamma_{k\uparrow} + u_k \gamma^\dagger_{-k\downarrow} \tag{20}$$

with $|u_k|^2 + |v_k|^2 = 1$.
If we insert these operators into the model Hamiltonian, use the relation between u_k and v_k, then properly choose u_k and v_k, the following relations will be obtained:

$$2\xi_k u_k v_k + \triangle^*_k v^2_k - \triangle_k u^2_k = 0 \tag{21}$$

Simplify above equation by multiple $\triangle^*_k / u^2_k$ to both two sides gives:

$$\frac{\triangle^*_k v_k}{u_k} = \sqrt{\xi^2_k + |\triangle_k|^2} - \xi_k \tag{22}$$

Excitation energy E_k is defined as $E_k = \sqrt{\xi^2_k + |\triangle_k|^2}$ [Fig. 5(a)], thus, v_k and u_k can be expressed in term of E_k as:

$$|v_k|^2 = 1 - |u_k|^2 = \frac{1}{2}(1 - \frac{\xi}{E_k}) \tag{23}$$

In BCS theory, Bardeen, Cooper and Schrieffer proposed the ground state as:

$$|\Psi_G\rangle = \prod_k (u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}) |0\rangle \tag{24}$$

Noted that $|u_k|^2 + |v_k|^2 = 1$. This implies that the parameter u_k and v_k are the probability amplitudes. The probability of the pair ($k\uparrow$,–$k\downarrow$) being occupied is $|v_k|^2$, while the probability of the pair being unoccupied is $|u_k|^2$. The relations between occupation probability $|v_k|^2$ and $|u_k|^2$ is shown in Fig. 5(b).
The ground state energy can be expressed as:

$$\langle E_G \rangle = \sum_k \left(\xi_k - \frac{\xi^2_k}{E_k} \right) - \frac{\Delta^2}{V_0} \tag{25}$$

Because the occupation probability is given by the Fermi-Dirac distribution, the definition of $\triangle_k$ can be rewritten as:

$$\begin{aligned} \triangle_k &= -\sum_{k'} V_{k,k'} \langle c_{-k\downarrow} c_{k\uparrow} \rangle \\ &= -\sum_{k'} V_{k,k'} u_{k'} v_{k'} (1 - 2f(E_{k'})) \\ &= -\sum_{k'} V_{k,k'} \frac{\triangle_{k'}}{2E_{k'}} tanh \left(\frac{E_{k'}}{2k_B T} \right) \end{aligned} \tag{26}$$

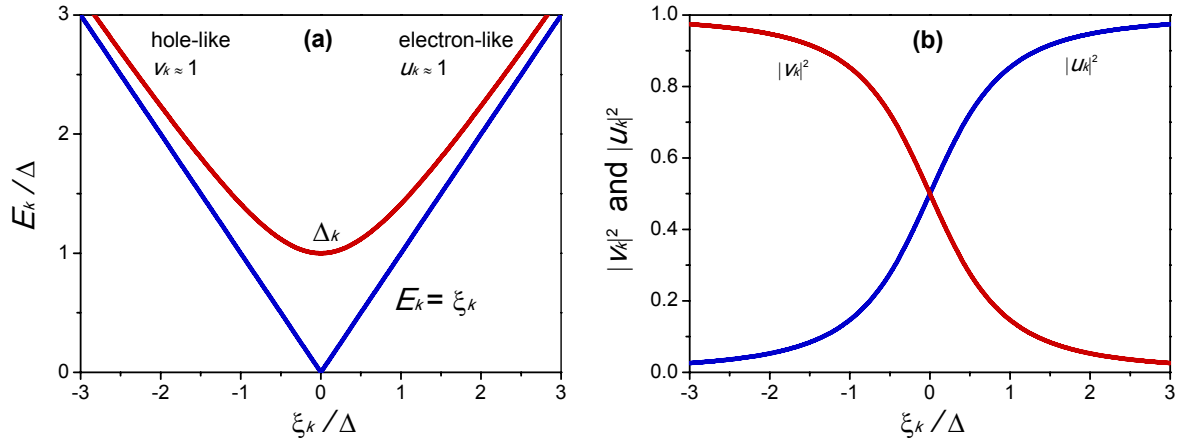


Fig. 5: *(a) Excitation energy E_k as a function of ξ_k in the normal and superconducting state. (b) Occupation probability $|v_k|^2$ and $|u_k|^2$ as a function of ξ_k at T = 0 near the Fermi level.*

Based on the assumption of week coupling in BCS theory, *i.e.* $V_{k,k'} = -V_0$, $\triangle_{k'} = \triangle$, the summation of above equation can be converted into an integration over energy, then we arrive at the BCS gap equation:

$$1 = \lambda \int_0^{\hbar\omega_D} \frac{1}{E_k} tanh\left(\frac{E_k}{2k_BT}\right) d\xi_k \tag{27}$$

where $\lambda = V_0 D(\mu)$ is the dimensionless electro-phonon coupling constant, $D(\mu)$ is the density of states. The BCS gap equation gives the temperature dependence of gap energy, in particular, it gives not only the energy gap at zero temperature but also the ordering temperature T_c.
For temperature approaching to T_c, we have $\triangle \to 0$. Then we can get the equation for the critical temperature:

$$k_BT_c = 1.13\hbar\omega_D e^{-1/\lambda} \tag{28}$$

Also at low temperature the ratio between $\triangle$ and T_c can be determined as

$$\frac{\triangle(0)}{k_BT_c} = 1.764 \tag{29}$$

Above relation is one of the most important result deduced from the BCS theory, indeed, it was obeyed by all classical metallic superconductors as an universal amplitude ratio.

3 A survey of the superconducting materials

After superconductivity was found in Hg with T_c = 4.2 K, great efforts are made to search for new superconductors. Till now, thousands of superconducting materials have been found [11]. However, superconductivity is still a low temperature phenomenon. To find a room-temperature superconductor seems still a long way to go. In the following parts, the main superconducting materials are classified.

• **Superconducting element**
After Hg, Superconductivity is also found in some other elements, such as Sn, Pb and La. Among all elements, Pb possesses the highest T_c of 9.2 K in the ambient pressure. Although some elements are non-superconductor at very low temperature, the superconductivity can emerge when they are subjected to high pressure or fabricated as thin films.

• **Superconducting alloys and compounds**
As a solid solution of different kinds of atoms, some alloys also exhibit superconductivity, such as NbTi (T_c = 9.5 K) and NbTa (T_c = 6.0 K). Higher T_c are also found in other Nb-content A_3B compounds, e.g. T_c = 18 K for Nb_3Sn and T_c = 23.2 K for Nb_3Ge. Besides, the superconductivity with T_c up to 40 K was discover in MgB_2 [12].

• **Organic superconductors**
Usually, the Organic compounds are insulators, but it was found that some organic compounds are superconductors. The Critical temperatures of organic superconductors are still in the range of classical superconductors [13].

• **Heavy-fermion superconductors**
In Heavy-fermion system, the electrons have large effective mass of about 200 times the free electron mass. Superconductivity was found in some heavy-fermion system such as $CeCu_2Si_2$. It was believed that heavy Fermion superconductors belong to the unconventional superconductor, however, the mechanism of this kind of superconductor is still not clear [14].

• **Superconducting fullerene compounds**
Fullerene, with the formula C_{60} was discovered as the third form of carbon in 1980s. After intercalating with exotic atoms, the doped C_{60} molecular crystals will exhibit the superconductivity with T_c up to 40 K [15, 16].

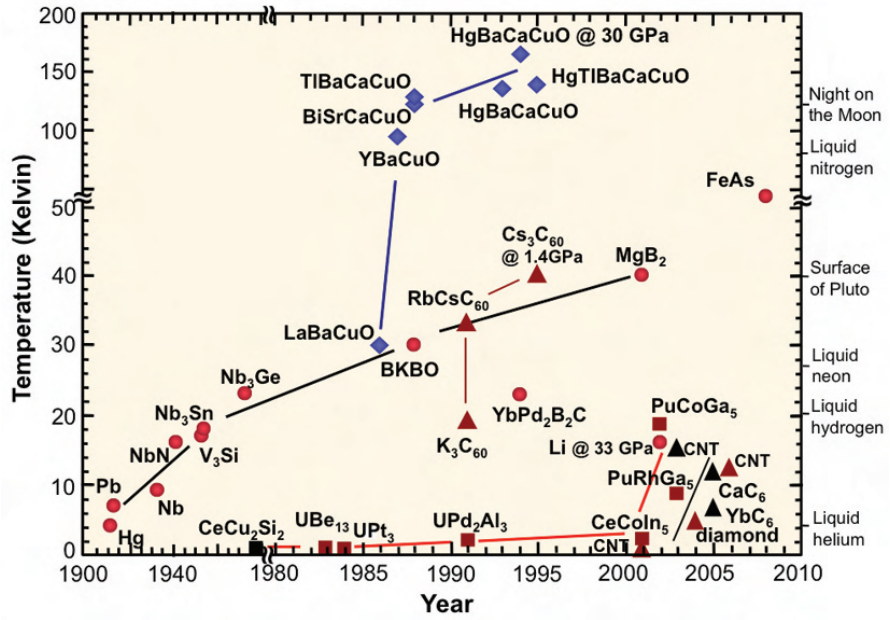


Fig. 6: *Timeline of the discovery of superconductors [17].*

• **High-T_c cuprate superconductors**

In 1986, the first cuprate superconductor $La_{2-x}Ba_xCuO_4$ with $T_c \approx 30$ K was discovered by Bednorz and Müller [2]. It was quite surprising that high T_c is shown in cuprate since the oxidize compounds are always insulators or poor conductors. The researchers are immediately motivated by this breakthrough and start searching for new cuprate compounds with higher T_c. Soon after, the superconductivity is observed in $YBa_2Cu_3O_{6+x}$ with $T_c \approx 90$ K, which is well above the temperature of liquid nitrogen [18]. So far the highest T_c of 133 K at ambient pressure was found in Hg-based cuprate $HgBa_2Ca_2Cu_3O_8$ [19]. Obviously, the Cuprate with high T_c are unconventional superconductors and it can not be explained in the framework of BCS theory. Therefore, new theory is required to explain the physics of the cuprate superconductors. However, the fundamental mechanism of high-temperature superconductivity is still unclear and it was still considered as the topic of the frontier of condensed matter physics.
From the point of view of crystal structure it was found that all high T_c cuprate have a layered structure. All Cuprate have one or more layers of copper oxide (CuO_2) and the CuO_2 layer are spaced by layers containing elements such as lanthanum, barium or yttrium. The schematic view of crystal structure of $YBa_2Cu_3O_7$ is shown in Fig. 7(a). Based on the large number of experimental results the universal phase diagram relating the critical temperature to doping level can be draw Fig. 7(b).

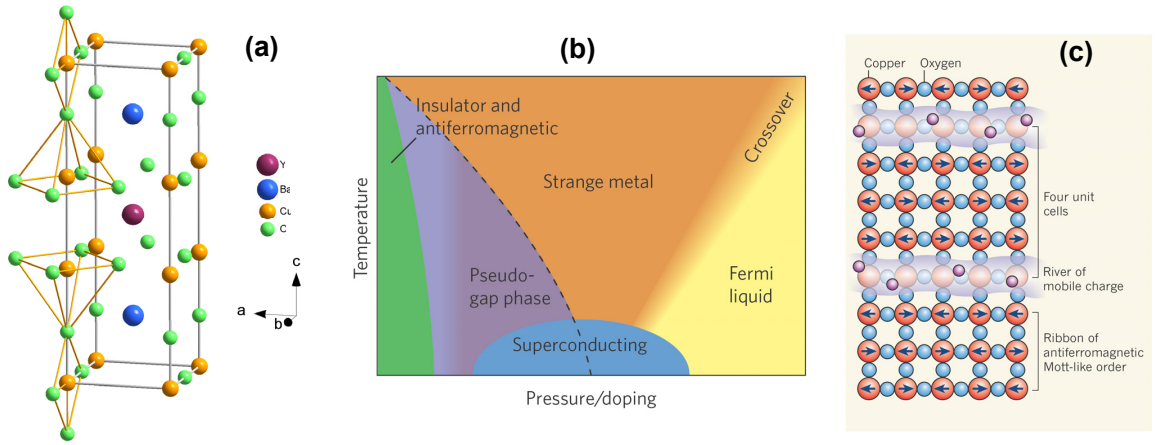


Fig. 7: *(a) Crystal structure of $YBa_2Cu_3O_7$. (b) Schematic phase diagram of hole doped cuprate high T_c cuprate superconductor [20]. (c) Stripe-like electronic order in the cuprates [21].*

Considering the phase diagram with increasing doping, the material is still an antiferromagnetic insulator at the lowest doping level. With increasing doping, the material enters the superconducting phase and the critical temperature T_c exhibits a dome-like dependence on doping with further increases in the doping level. At the over doped level, the material again become non-superconducting. Between the antiferromagnetic and superconducting phase zone, there is phase named pseudogap phase, in which physical properties show behavior of the existence of an energy gap. It is still controversial whether the pseudogap arisen from competing orders or it is the precursor of superconductivity.
As shown in Fig. 7(b), the undoped Cuprate compounds are Mott insulator with long-range antiferromagnetic order of Cu. Upon doping with holes, the stripes of spin and charge order formed. Both charge and spins are periodically modulated in the stripe phase [22]. As illustrated in Fig. 7(c), inhomogeneity arises due to the hole doping process. The antiferromagnetic spin order in the spin-part of stripe is similar as in the undoped antiferromagnetic Mott insulator, whereas the charge can conduct between the spin-part of stripe. Experimentally, the

strip phase is directly detected by neutron scattering study. Information on the period of both charge and spin density modulations can be obtained. It was also observed that the period of the charge order is temperature dependent, which indicates the variation of charge concentration as a function of temperature. Theoretically, it was thought that the formation of stripe phase was attributed to the competition between the kinetic energy of the electrons, the antiferromagnetic interaction among spins, and the Coulomb interaction between charges. It was believed that the existence of the strip phase might give rise to superconductivity.

• **Iron-based high-T_c superconductors**

The iron-based superconductors, discovered in 2008 [3], is the latest family of high-T_c superconductors. Extensive attention are drawn afterwards since it provides a new opportunity to investigate the mechanism of unconventional superconductivity. More detailed introductions on iron-based superconductors can be found in the next section.

4 Iron-based high-T_c superconductors

4.1 Materials series and phase diagrams

In 2008, a new family of high-T_c superconductors, iron-based superconductors were discovered [3]. This discovery has provided a new playground to investigate the mechanism of unconventional superconductivity [23]. There are a number of homologous families of iron-based superconductors discovered so far, which are short-named after the stoichiometries of their parent compounds (as shown in Fig. 8(a)). While most of the iron-based superconductors contain arsenic, some contain phosphorus, which come from the same pnictogen group of the periodic table. Therefore, they are usually termed as iron pnictides. In the so-called "11" and "245" families, the pnictogen is replaced by selenium or tellurium from the chalcogen group. These families are thus termed as iron chalcogenides. All iron-based superconductors have a common layer of iron atoms which are tetrahedrally coordinated by pnictogen or chalcogen atoms (as shown in Fig. 8(c)). They differ only in the details of the buffer layers. Experiments and theory now agree that the superconducting electrons in all the iron-based superconductors flow in the planes that contain Fe. Despite of the compositional variety, the Fe-containing planes have the same structure from material to material (as shown in Fig. 8(b)). Note that the iron-containing plane is not flat: pnictogen or chalcogen atoms reside above and below the plane. Because the pnictogen and chalcogen atoms are much larger than the iron atoms, they pack themselves in edge-sharing tetrahedral. By contrast, the smaller size difference between the copper and oxygen atoms in a cuprate superconductor leads to corner-sharing octahedral packing. Such a structural difference has an important consequence on the respective crystal-field splitting and electronic structure.

Most of the research on iron-based superconductors has focused on RFeAs($O_{1-x}F_x$)(with R = La, Nd, Sm, or Pr etc.) and AFe$_2$As$_2$ (with A = Ba, Ca, or Sr etc.), the so-called "1111" and "112" families. Starting from the parent compounds, superconductivity can be achieved either by doping, or by the application of pressure in some materials. Since the parent compounds are already metallic, the effect of the doping can not be solely related to the introduction of free charge carriers. It has been suggested that the modifications of the Fermi surface, which are similar under pressure and chemical doping, are important for inducing superconductivity

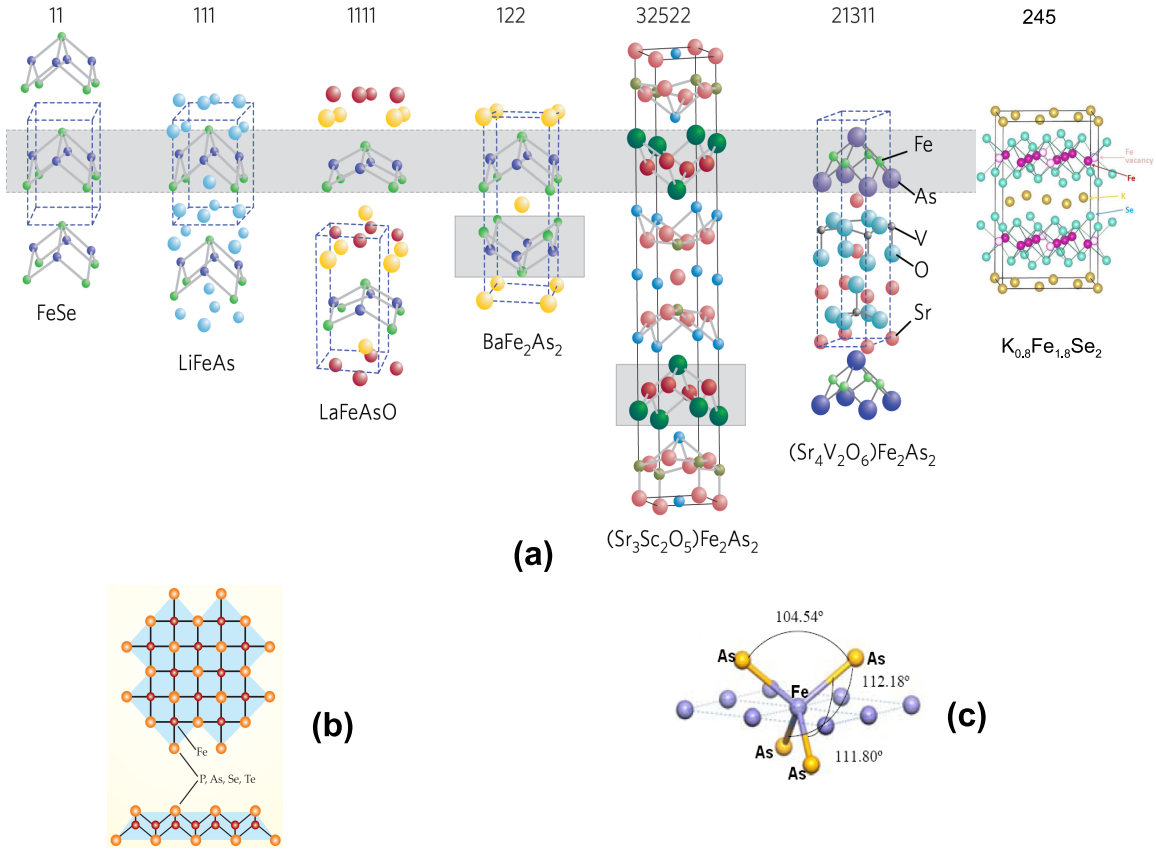


Fig. 8: *(a) Several homologous families of iron-based superconductors [23]; (b) the crystalline plane containing of iron pnictides or iron chalcogenides; (c) the iron atom is tetrahedrally coordinated by pnictogen or chalcogen atoms.*

in the iron-based compounds. Up to now, the highest T_c attained is 57.4 K in the electron-doped '1111' compound $Ca_{0.4}Nd_{0.6}FeAsF$, while for the '122' compound the highest T_c of 39 K is reached in the hole-doped $Ba_{0.6}K_{0.4}Fe_2As_2$. A typical behaviour of the resistivity and magnetization as a function of temperature of the electron-doped $Ba(Fe_{1-x}Co_x))_2As_2$ single crystals is shown in Fig. 9(a)-(b) [24]. The resistivity anomaly at 137 K in the undoped parent 122 compound is due to concurrent phase transitions from tetragonal to orthorhombic and from Pauli paramagnet to low-temperature spin-density wave (SDW) phase. With the increasing of the Co doping level, both orthorhombic structural and SDW phases are suppressed, while superconductivity emerges eventually at higher dpoing levels.

The measurements of resistivity, magnetic susceptibility as well as more bulk-sensitive techniques, such as thermal expansion, heat capacity and neutron and x-ray diffraction etc. would allow for an accurate determination of the phase diagrams of the iron-based superconductors. The phase diagrams of three representative iron-based superconductors are shown in Fig. 10. One of the most fascinating phenomena is the apparent coexistence and competition between superconductivity and the SDW phase in the underdoped regime of $Ba(Fe_{1-x}Co_x)_2As_2$. The transition from the SDW phase to superconductivity in $LaFeAs(O_{1-x}F_x)$ appears more abrupt. Such coexistence between superconductivity and magnetism has also been observed in the hole-doped $Ba_{1-x}K_xFe_2As_2$. The exact nature of this phenomenon remains to be established.

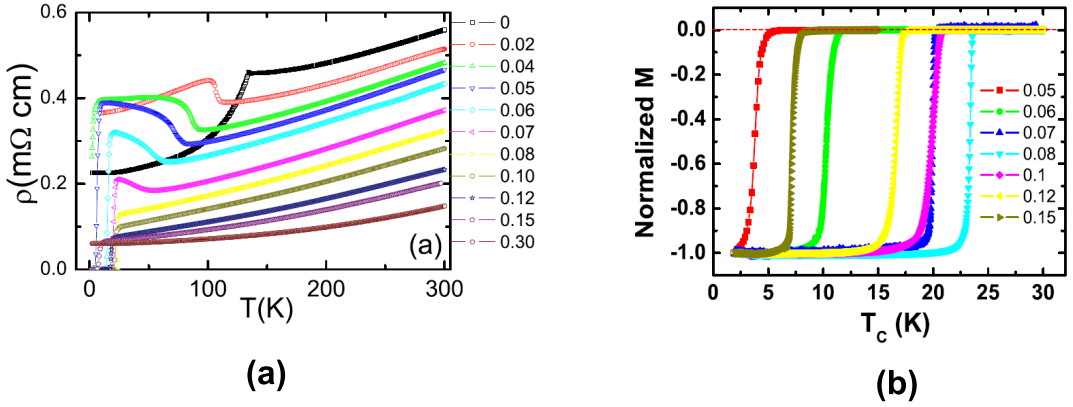


Fig. 9: *(a) The temperature dependence of (a) the resistivity and (b) the magnetization of the electron-doped $Ba(Fe_{1-x}Co_x)_2As_2$ single crystals [24].*

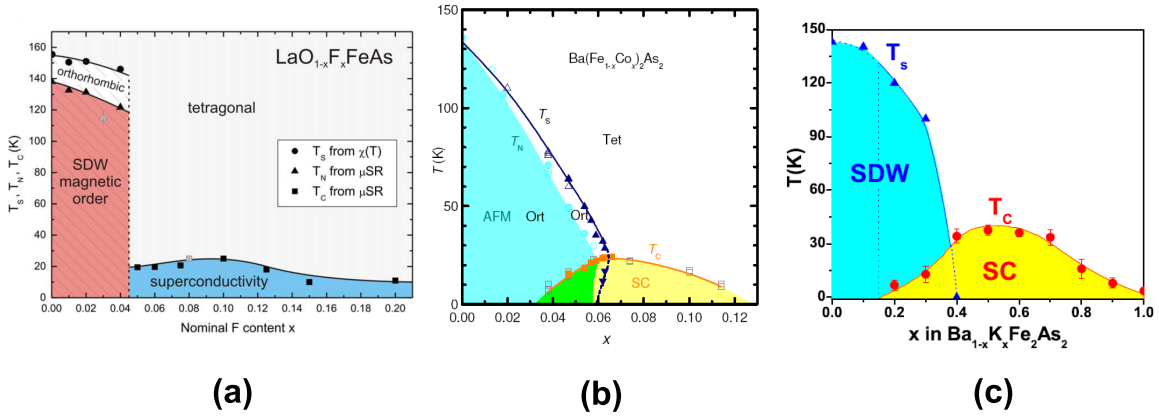


Fig. 10: *(The phase diagram of: (a) $LaFeAs(O_{1-x}F_x)$[23]; (b) Co-doped $Ba(Fe_{1-x}Co_x)_2As_2$ [25]; (c) $Ba_{1-x}K_xFe_2As_2$. [23]*

4.2 Close proximity to magnetism: magnetic ordering and spin fluctuations

The nature of the magnetic ordering and spin fluctuations in superconductors has had a rich and interesting history, and has been a topic of special interest ever since the parent compounds of the high-T_c cuprates were found to be antiferromagnetic Mott insulators that exhibit huge exchange energies within the Cu-O planes. These strongly correlated spin fluctuations persist into the superconducting regime, often developing a spin resonance mode whose energy scales with T_c and whose intensity exhibits a superconducting order-parameter-like behavior. Recently discovered iron-based superconductors represent another remarkable example in which superconductivity is in close proximity to magnetism. Although neutrons do not couple directly to the superconducting order parameter, they have nevertheless played a decisive role in the understanding of the interplay between superconductivity and magnetism, as demonstrated by the determination of magnetic ordering in the parent compound and the observations of the spin resonant mode in the superconducting counterparts. It has also been established via neutron

scattering that magnetic ordering in iron pnictides is strongly coupled to lattice instabilities.

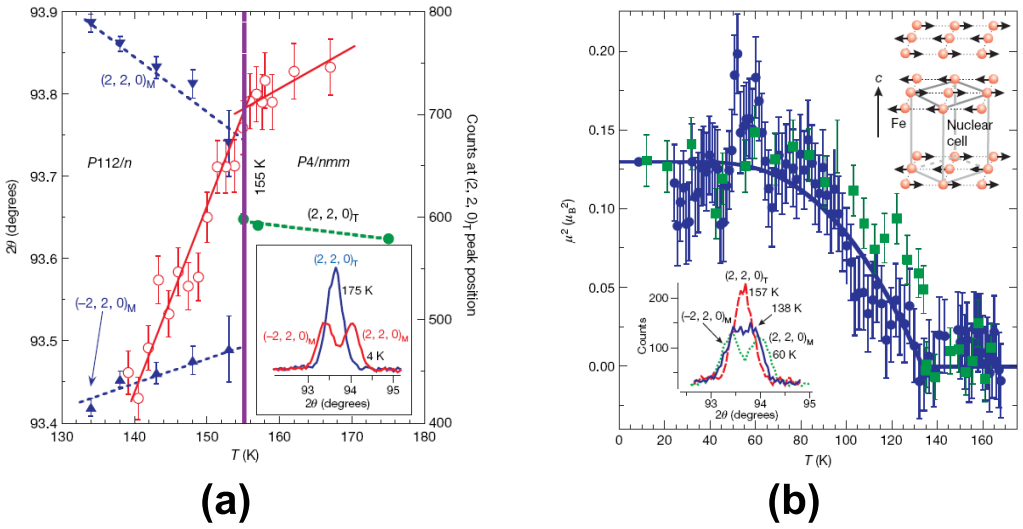


Fig. 11: *Powder neutron diffraction of LaFeAsO [26] : (a) a tetragonal-to-orthorhombic transition occurs at 155 K; (b) an antiferromagnetic ordering due to SDW occurs at 138 K.*

As shown in Fig. 11, the coupling of the antiferromagnetic and structural phase transitions in LaFeAsO has been captured via powder neutron diffraction [26]. This allowed for the determination of long-range collinear antiferromagnetic ordering with a very small saturation moment of 0.4 μB per Fe site.

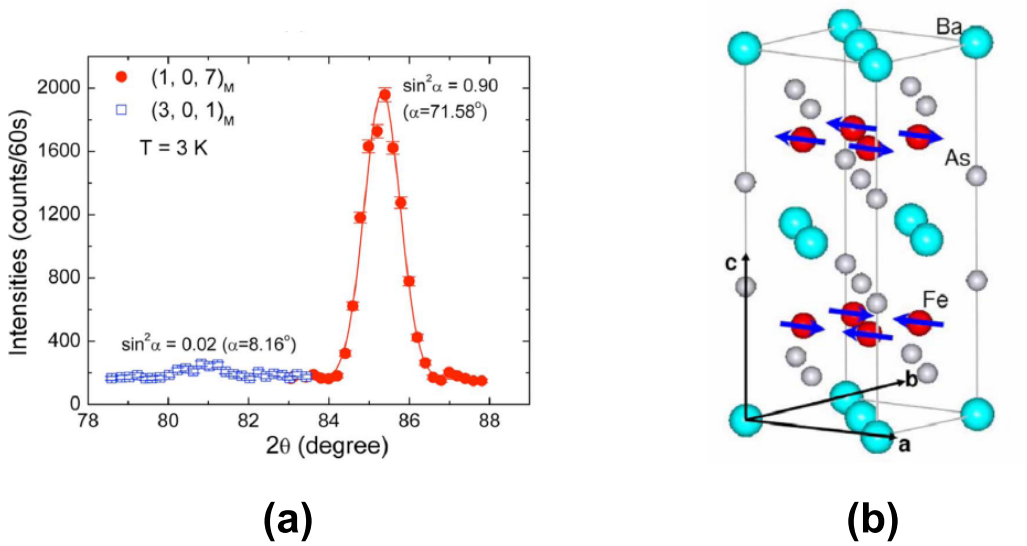


Fig. 12: *Neutron diffraction of single-crystalline $BaFe_2As_2$: (a) the measurement of two magnetic reflections, the clues about the aligning direction of the ordered magnetic moment can be obtained from the drastically different peak intensities; (b) the determined magnetic structure of $BaFe_2As_2$ [27].*

Such a collinear antiferromagnetic ordering was soon after confirmed in single-crystalline $BaFe_2As_2$ [27] and other iron pnictide compounds [23]. As shown in Fig. 12(a), from the measurement

of these two magnetic reflections with a modulation wavevector (1, 0, 1), the clues about the aligning direction of the ordered magnetic moment can be obtained from the drastically different peak intensities. The ordered magnetic moment of iron has been determined to be aligned along the longer a-axis of the low temperature orthorhombic phase. The magnetic moments of iron are arranged in such a way that they are anti-parallel to the neighbouring ones along the orthorhombic a- and c-axis, while they are parallel along b-axis (in Fig. 12(b)). Except the case in the parent compound of iron chalcogenides family, the saturation moments of Fe in all other iron pnictide parent compounds have been found to be in the range of 0.3-1.0 μB.

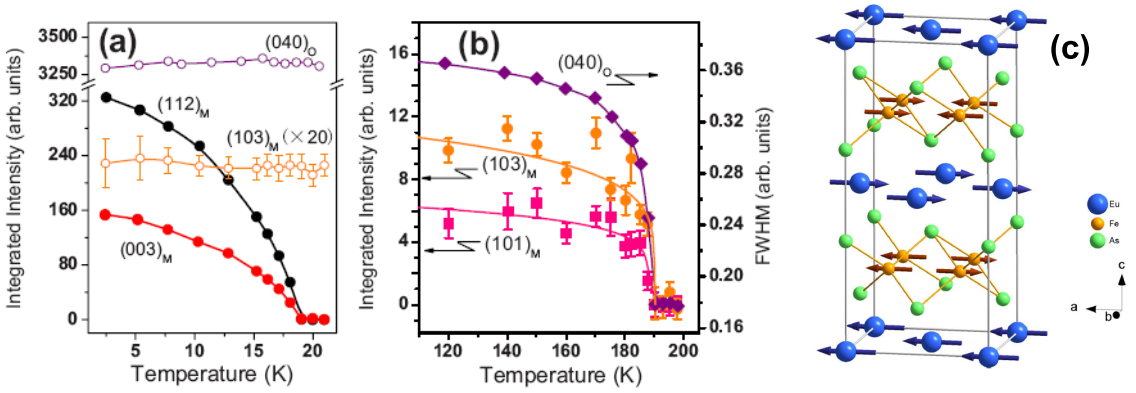


Fig. 13: *The magnetic structure of $EuFe_2As_2$: (a)-(b) the temperature dependence of the Eu^{2+} antiferromagnetic reflections $(112)_M$ and $(003)_M$ and the Fe-SDW reflections $(101)_M$ and $(103)_M$; (c) the refined magnetic structure [28].*

However, the determination of magnetic structures in some magnetic rare-earth containing iron pnictide compounds is not trivial due to the presence of two magnetic sublattices as well as the possible interplay between the localized rare-earth magnetism and itinerant Fe-SDW, such as in $EuFe_2As_2$ [28] and SmFeAsO [29] etc. Via complementary single-crystal neutron and magnetic x-ray scattering, the magnetic structure of $EuFe_2As_2$ has been thoroughly determined [28](see Fig. 13). The Fe-SDW retains the same magnetic ordering as in other parent iron pnictides, the the Eu^{2+} moment is aligned along the orthorhombic **a**-axis. The propagation wavevector of the Eu^{2+} antiferromagnetic ordering is $\mathbf{Q}_{Eu}$ = (0, 0, 1) with T_N = 19 K, while the Fe-SDW orders at $\mathbf{Q}_{SDW}$ = (1, 0, 1) with T_{SDW} = 190 K. The temperature dependence of the corresponding magnetic reflections indicates a rather weak coupling between these magnetic sublattices. The situation appears quite differently in SmFeAsO, in which a strong coupling between Sm and Fe magnetism has been experimentally demonstrated [29].

While the static magnetic ordering in various iron pnictides compounds has been well established, the nature of magnetism is being hotly debated. Two approaches have been used so far, one based on the localized Heisenberg J_1-J_2 exchange interactions, the other based on the itinerant picture where magnetism is governed by the Fermi surface (FS) nesting wavevector between the hole pockets at the Γ-point and the electron pockets at the M-point (as schematically shown in Fig. 14). A recent inelastic neutron scattering investigation on the spin-wave excitations of $CaFe_2As_2$ [30] indicated that the spin waves in the entire Brillouin zone can be described by an effective three-dimensional local-moment Heisenberg Hamiltonian, but the large in-plane anisotropy cannot. Therefore, magnetism in the parent compounds of iron arsenide superconductors was suggested to be neither purely local nor purely itinerant, rather it is a complicated

mixture of the two.

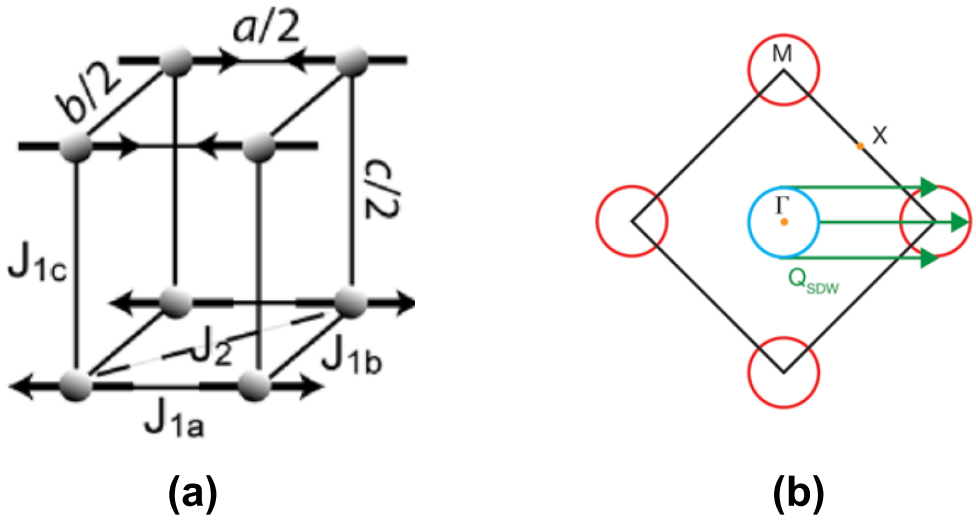


Fig. 14: *The nature of magnetism in iron-based superconductors: (a) Heisenberg J_1-J_2 exchange interactions based on the well localized spins; (b) Interband particle-hole excitation due to the Fermi surface nesting.*

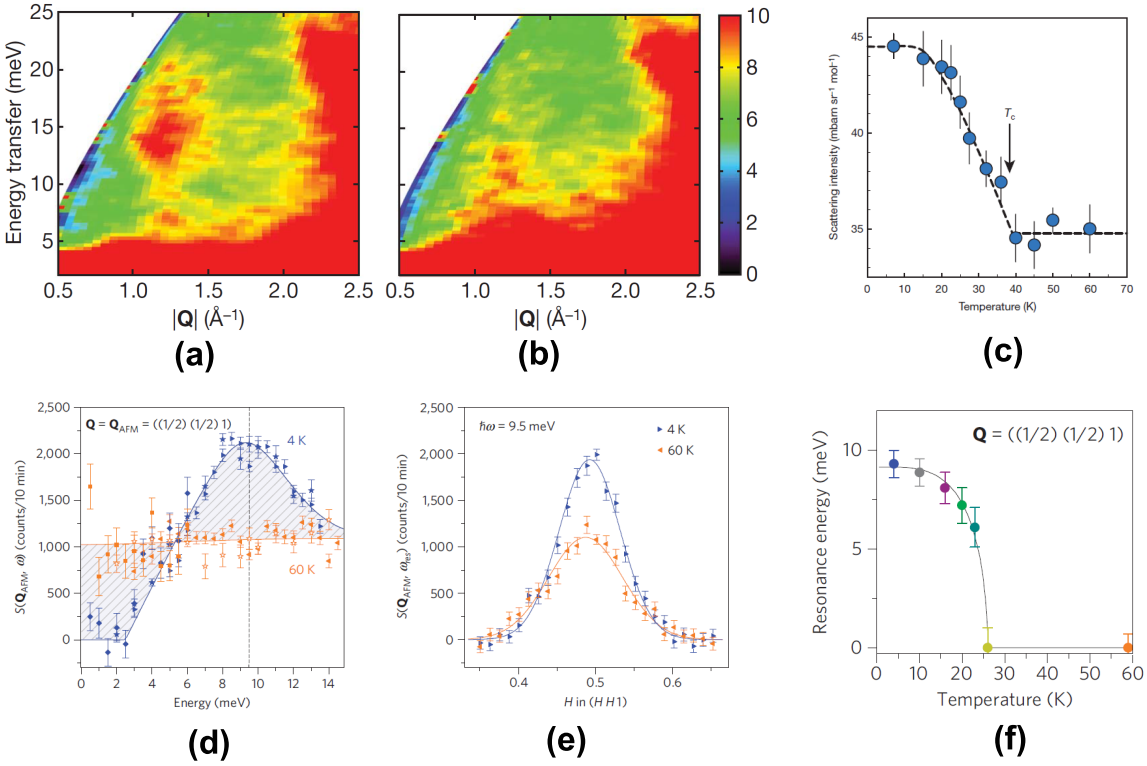


Fig. 15: *The spin resonance mode observed via inelastic neutron scattering techniques: (a-c) in $Ba_{1-x}K_xFe_2As_2$ via time-of-flight spectroscopy [31]; (d-f) in $Ba(Fe_{1-x}Co_x)_2As_2$ via triple-axis spectroscopy [32].*

Measurements of the spin dynamics within the spin density wave phase of the parent compounds has shown evidence of strongly dispersive spin waves with exchange interactions consistent with the observed magnetic order and a large anisotropy gap. Again, in a way very similar to the cuprates, these antiferromagnetic spin fluctuations persist in the normal state of the superconducting compounds, but they become more diffuse. Below T_c, there exists evidence in several "122" compounds that these spin fluctuations condense into a resonant spin excitation at the in-plane antiferromagnetic wavevector with an energy that scales with T_c (see Fig. 15) [31, 32, 23] . Such resonances have been also observed in the high-T_c cuprates and a number of heavy fermion superconductors, where they are considered to be the evidence of d-wave symmetry. Since it has been indicated from ARPES and other measurements that the superconducting gap in iron-based superconductors is likely isotropic, the observation of the spin resonance mode in neutron scattering has thus been seen as strong evidence of unconventional superconductivity due to sign-reversal $s_{\pm}$-wave symmetry.

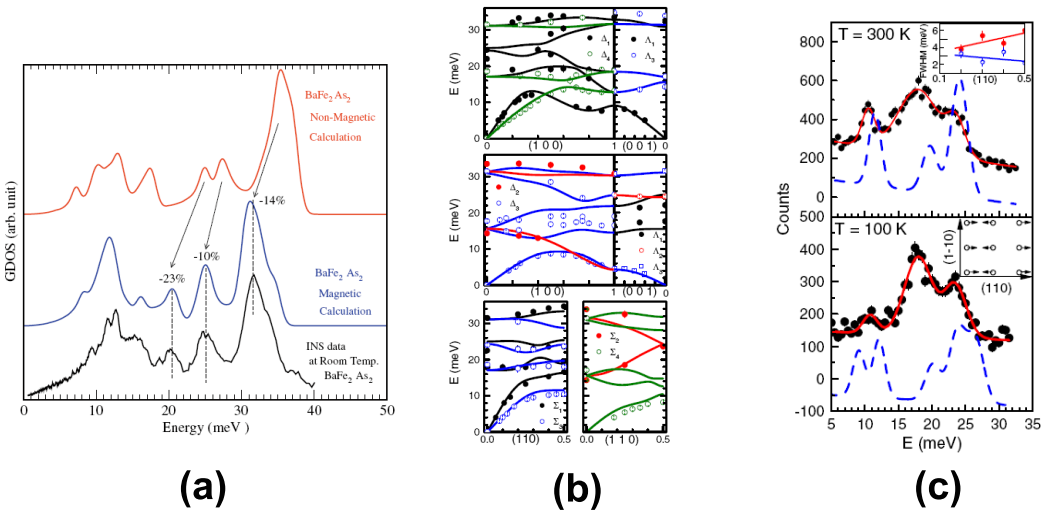


Fig. 16: *(a) Generalized phonon-DOS measured by inelastic neutron scattering at room temperature for $BaFe_2As_2$ (bottom black curve) and the calculated GDOS with (middle blue) and without (top red) Fe-magnetism [33]; (b) comparison of experimentally determined phonon frequencies (solid circles) in the (100), (001) and (110) directions of $CaFe_2As_2$ at T = 300 K with the results of density functional theory (solid lines)[34]; (c) anomalous phonons in $CaFe_2As_2$ measured at Q = (2.5, 1.5, 0) at room temperature and at a temperature far below the structural phase transition. The calculated phonon structure factors for nonmagnetic and spin-polarized are shown in the upper and lower panel (dashed lines), respectively [34].*

4.3 Roles of electron-phonon coupling

It has been widely accepted that electron-phonon coupling alone can not account for the occurrence of high temperature superconductivity in iron-based superconductors [35]. However, the importance of electron-phonon coupling have been implied from a number of experiments. As shown in Fig. 16(a), the generalized phonon density-of-state (DOS) measured by inelastic neutron scattering at room temperature for $BaFe_2As_2$ is much closer to the calculated magnetic phonon-DOS rather than non-magnetic one [33]. This clearly demonstrates that the Fe-magnetism is intimately coupled to the phonons in iron pnictides. One way to verify possible

roles of electron-phonon coupling is to measure phonon dispersions as a function of temperature in details on single-crystal sample via inelastic neutron scattering. A fairly complete picture of phonon dispersions in the main symmetry directions of $CaFe_2As_2$ has been established from such an approach [34], as shown in Fig. 16(b). Strong temperature dependence of some phonons near the structural phase transition near 172 K was observed (see Fig. 16(c)). The calculated phonon spectra for non-magnetic/spin-polarized structures are shown as dashed lines in Fig. 16(c). The agreement between the experimental results and the calculation is poor. It has been suggested that the interplay between magnetism and the lattice is in some way responsible for the anomalous phonons in $CaFe_2As_2$. That is to say, the coupling of the vibrational and the electronic degrees of freedom is stronger than calculated by density functional theory, and hence phonon might play an important role in superconductivity in the doped compounds.

4.4 Superconducting gap symmetry and structure

Understanding the nature of the superconducting gap (Δ) in a superconductor is essential to establish the microscopic origin of superconductivity, since the magnitude, symmetry and structure of a superconducting gap are directly associated with the pairing strength and the pairing interactions. For instance, the wider the gap, the harder it is to break apart the Cooper pairs and destroy the superconducting state. In conventional superconductors (e.g. Nb, Pb and Al etc.), the superconducting gap function, that can be nicely predicetd by the BCS theory, has *s*-wave symmetry and it is thus isotropic along the Fermi sea (as schematically shown in Fig. 17(a)). This reflects the pairing interaction being attractive, its origin is the electron-electron interaction mediated by phonons. *p*-wave pairing symmetry has been suggested for superfluid ^{3}He and possibly in ruthenates superconductors. The pairing symmetry for the high-TC cuprates has been identified to be *d*-wave. The gap nodes, where the superconducting gap can be closed, exist for both *p*-wave and *d*-wave (see Fig. 17(b)-(c)).

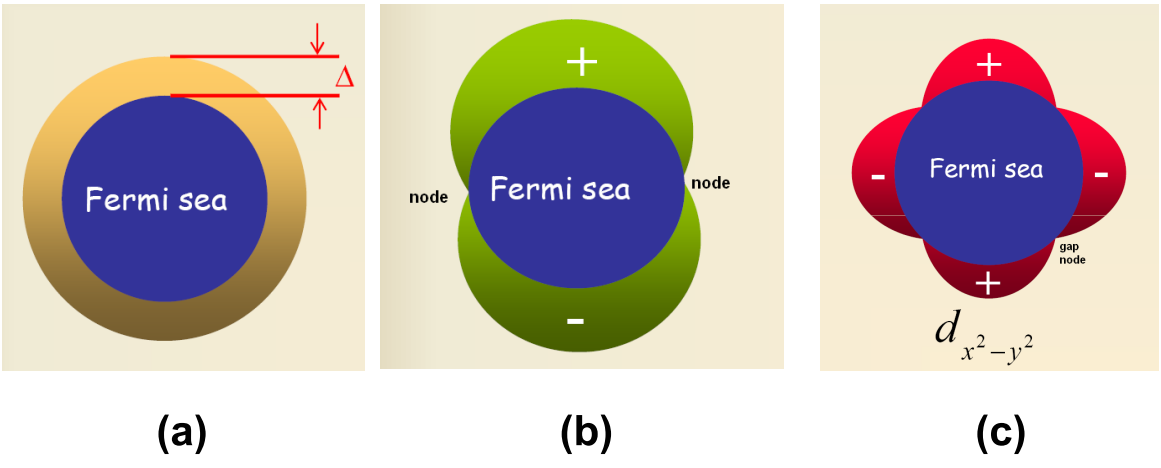


Fig. 17: *The schematic drawing of the superconducting gap functions: (a)* s*-wave, nodeless isotropic gap; (b)* p*-wave, with nodes; (c)* d*-wave, with nodes.*

One of the key challenges for iron-based superconductors is to identify the superconducting gap symmetry and structure. While the leading candidate for the pairing symmetry in iron-based

superconductors is a sign-reversal $s_{\pm}$-wave as first proposed by Mazin[36], unlike the case in cuprates, an unambiguous picture has not been established [23]. It has been found that the superconducting gap seemingly shows strong material dependence. As an example shown in Fig. 18, two isotropic gaps can be identified in hole-doped $Ba_{0.6}K_{0.4}Fe_2As_2$ via ARPES [37].

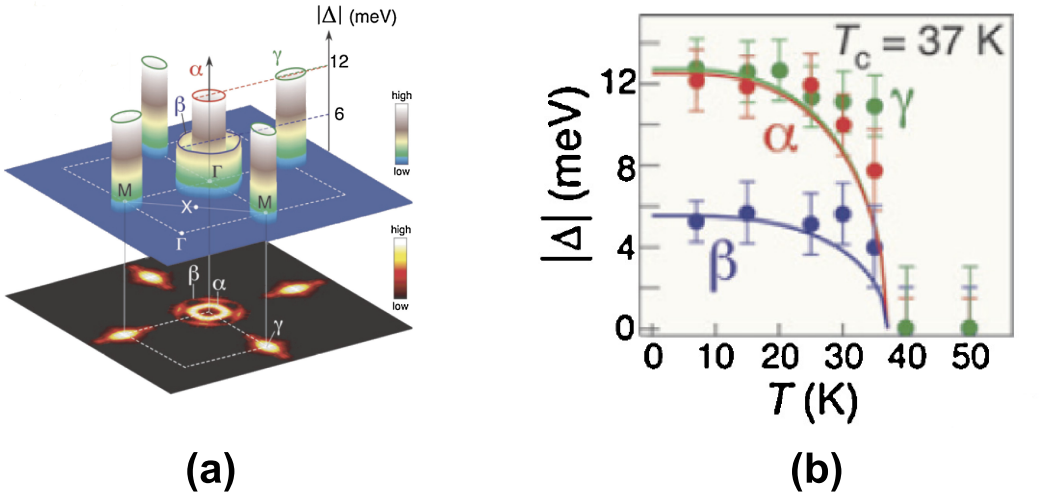


Fig. 18: *Fermi-surface-dependent nodeless superconducting gaps in $Ba_{0.6}K_{0.4}Fe_2As_2$: (a) three-dimensional plot of the superconducting gap amplitude (Δ) measured at 15 K on the three observed Fermi surface sheets (shown at the bottom as an intensity plot); (b) their temperature evolutions [37] .*

5 Special topics

5.1 Superconducting vortex lattices

For type-II superconductors with $\lambda \geqslant 1/\sqrt{2}\,\xi$, the magnetic flux can penetrate the superconductor and form the vortices. The superconducting order parameter will be suppressed over the core region of the length scan ξ and the supercurrent surrounding the vortex core will extend over the length scale λ. To minimize the total free energy of the system, these vortices will arrange themselves in a periodic array, which is called Abrikosov flux line lattice (FLL). Many important information concerning the superconducting state, such as penetration depth and coherence length can be obtained by investigating the flux line lattice [38].

Among many experimental techniques, Small Angle Neutron Scattering (SANS) is considered as one of the most suitable probe for flux line lattice because neutrons can be scattered via the interaction of their intrinsic moments with the modulation of magnetic field originated from the flux line lattice. The diffraction pattern from SANS provides not only the information about the structure of flux line lattice and its correlation with the crystal lattice, but also the information on the gap structure and the value of characteristic length scales [39].

In Fig. 19, series of SANS diffraction patterns from the flux line lattice of $YBa_2Cu_3O_{7-\delta}$ cuprate superconductor are presented [40]. these patterns are collected at 2 K and various magnetic field along the crystallographic *c* axis. It can be seen that hexagonal FLL structure is formed at 1.5

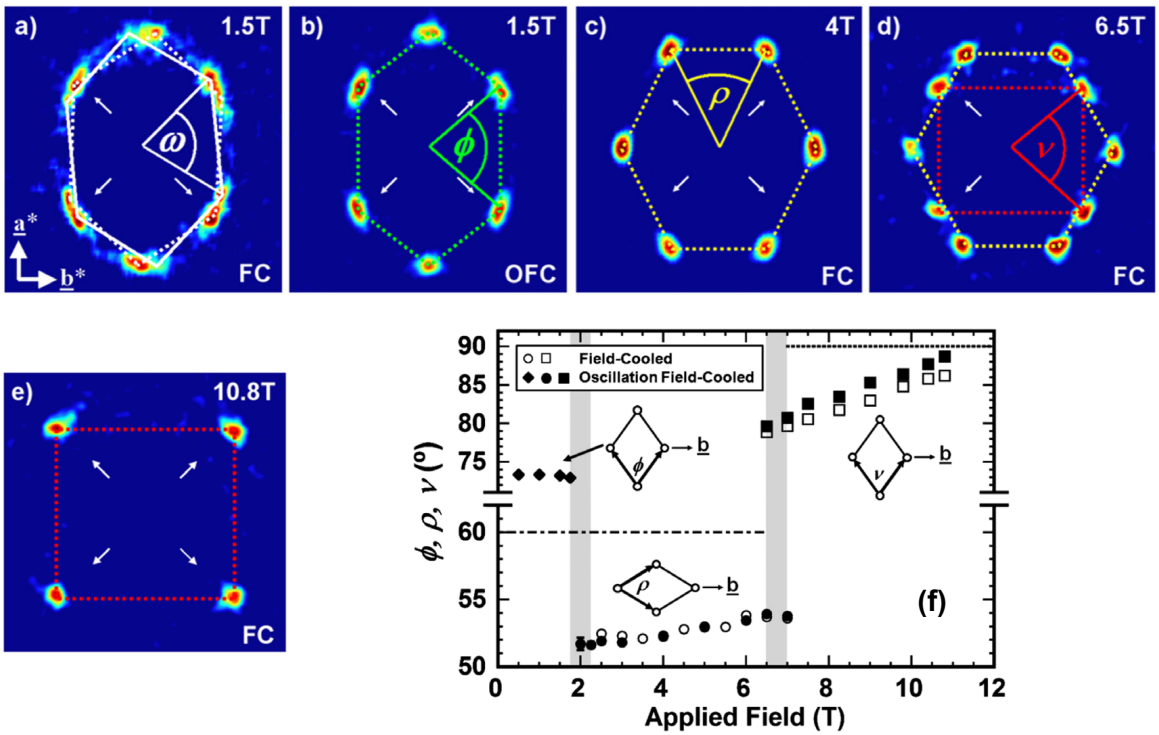


Fig. 19: *(a) to (e) Flux line lattice diffraction patterns of $YBa_2Cu_3O_{7-\delta}$ cuprate superconductor taken at 2 K in different magnetic fields. FC and OFC indicate field cooling and oscillation field cooling procedures, respectively. (f) Magnetic field dependence of the flux line lattice apex angle for different structure type [40].*

T [Fig. 19(a) and (b)]. When 4 T field is applied, a single distorted hexagonal FLL is observed. The transition between these two hexagonal FLL is likely to be first order as indicated by sudden change of FLL symmetry angle [see Fig. 19(f)]. Another first order FLL structure transition is observed with further increase of field. As shown in Fig. 19(d), the rhombic FLL structure emerges and coexists with the distorted hexagonal structure at 6.5 T. Once field reached above 7 T, the pure rhombic structure is established for the FLL.

In the simplest approximation, hexagonal flux line lattice is the most stable state due to the close packed structure. However, repulsive interaction between the flux line, property of the Fermi surface, as well as the anisotropic superconducting order parameter can result in the transition from hexagonal to other symmetry. As described above, two field driven first order flux line lattice structure transitions was observed in $YBa_2Cu_3O_{7-\delta}$. It is suggested that the low field transition is probably driven by Fermi surface effects, while the high field transition from hexagonal to square structure is due to the dominant role of anisotropic superconducting order parameter.

5.2 Crystal field excitation in superconductors

For the rare earth ions, both electric field and magnetic field can lift, at least partially lift the $2J+1$ fold degenerate ground state. Considering a three dimentional crystal structure, the electric field might be generated by the ion surrounding the rare earth ion in lattice. Following some definitions and transformations, crystal electric field (CEF) Hamiltonian can be expressed

as this famous notation, $\mathcal{H}_{CEF} = \sum_{l,m} B_l^m O_l^m$, where B_l^m is the crystal field parameters and O_l^m is the Stevens equivalent operators. It is obvious that the crystal field Hamiltonian will give rise to discrete energy levels. Neutron scattering is the method of choice to determine crystal field level schemes by measuring the excitation between these levels. The magnetic scattering of the CEF transition can be expressed in terms of differential neutron cross section. Because CEF of rare earth is largely related to the surrounding environment, to study the CEF of rare earth in rare earth based superconductor can provide detailed information on the local charge distribution and the formation of energy gap.

Inelastic neutron scattering has been long time used to investigate the relaxation behavior of the ground state crystal field excitation associated with the rare earth ion in high T_c cuprate superconductor. Since the crystal field levels are subject to an interaction with charge carriers, neutron scattering measurements of the relaxation rate of crystal field excitations can provide the direct evidence on the opening of energy gap and reflect the variation of density of state at the Fermi energy. For example, the pseudogap opening temperature in Ho-based cuprate can be clearly observed by following the emperature dependence of the intrinsic linewidth corresponding to the lowest ground state crystal field excitation [41].

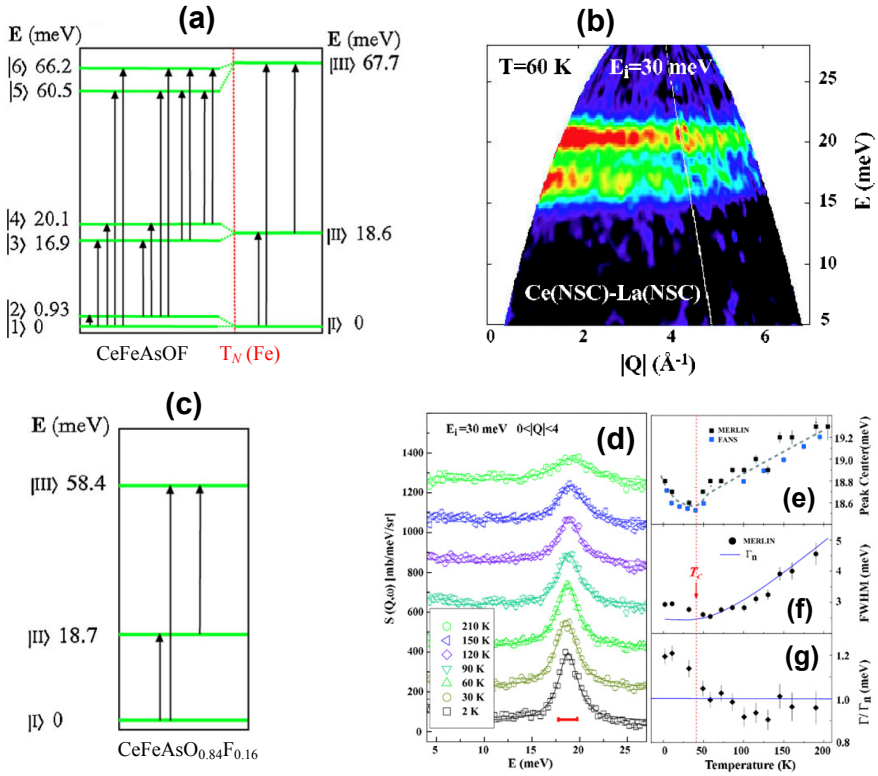


Fig. 20: *(a) Energy level scheme of Ce^{3+} ion in CeFeAsO as determined by inelastic neutron scattering. (b)Energy spectra of neutron scattered from CeFeAsO at T = 60 K. Noted that the phonons contribution collected using isostructural LaFeAsO compound was substracted. (c) Energy level scheme of Ce^{3+} ion in $CeFeAsO_{0.84}F_{0.16}$ as determined by inelastic neutron scattering. (d) Temperature dependence of excitation between ground state and the first excited state in $CeFeAsO_{0.84}F_{0.16}$. (e) Temperature dependence of the peak position (e), intrinsic line width (f), line width ratio (g) for the excitation peak in (d) [42].*

Regarding to the recently discovered Fe-pnictide superconductor, there also exist a large number of compounds that content rare earth elements, such as *R*FeAs($O_{1-x}F_x$) compounds with *R* = Ce, Nd, Sm, or Pr etc. As we will show in the following example, the crystal electric field can be used as a probe for long range antiferromagnetic order and superconducting state in $CeFeAsO_{1-x}F_x$ superconductors [42].
Similar to other parent compounds of Fe-pnictides, CeFeAsO exhibits a phase transition from tetragonal to orthorhombic structure and then orders antiferromagneticly with decreasing temperature. The crystal field Hamiltonian for two different structures can be written as $\mathcal{H}_{(T)} = B_2^0O_2^0 + B_4^0O_4^0 + B_4^4O_4^4$ and , $\mathcal{H}_{(O)} = B_2^0O_2^0 + B_2^2B_2^2 + B_4^0O_4^0 + B_4^2O_4^2 + B_4^4O_4^4$, respectively. The crystal field parameters B_l^m can be deduced by modeling the inelastic neutron scattering data. It was found that three doublets are shown in paramagnetic phase of CeFeAsF, and these three doublets split into six singlet when the Fe ions order antiferromagneticlly. The splitting scheme of crystal electric field is shown in Fig. 20(a). A typical inelastic neutron spectra of CeFeAsF collected at 60 K is presented in Fig. 20(b), in which two clear bands of CEF excitations at 16.9 and 20.1 meV are clearly observed.
For F-doped $CeFeAsO_{0.84}F_{0.16}$ superconductor, the tetragonal structure maintains in wide temperature range from 4 K to room temperature. Three doublets are detected [Fig. 20(c)]. In Fig. 20(d), the excitation between the ground state and the first excited state is plot as a function of temperature. The temperature dependence of peak position, intrinsic line width are obtained by fitting the excitation peaks [Fig. 20(e),(f) and (g)]. Both peak position and line width exhibit anomalies at critical temperature T_c at around 40 K, which suggested that CEF can be used as a probe for the superconducting state in Fe-pnictide superconductors.

6 Summary

While high temperature superconductivity remains one of the biggest challenges in condensed matter physics, the understanding of its mechanism has advanced tremendously over the past 26 years or so. A common feature among cuprates, iron-based superconductos and heavy-fermion superconductors is that superconductivity is in close proximity to magnetism. This has thus provided neutron scattering with an important playground due to its unique sensitivity to magnetic correlations in a wide dynamic range as demonstrated in previous sections. It can be expected that the advanced scattering methods in particular neutron scattering will continue to play a major role in the studies of superconductivity.

References

[1] H. K. Onnes, Leiden Commun. **120b, 122b, 124c** (1911).

[2] J. G. Bednorz and K. A. Müller, Zeitschrift für Physik B Condensed Matter **64**, 189 (1986).

[3] Y. Kamihara *et al.*, J. Am. Chem. Soc. **130**, 3296 (2008).

[4] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc. New York, 2004).

[5] P. Hofmann, *Solid State Physics* (Wiley-VCH, Berlin, 2008).

[6] W. Meissner and R. Ochsenfeld, Naturwissenschaften **21**, 787 (1933).

[7] F. London and H. London, Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences **149**, 71 (1935).

[8] V. L. Ginzburg and L. D. Landau, Zh. Eksp. Teor. Fiz. **20**, 1064 (1950).

[9] A. A. Abrikosov, Sov. Phys. JETP **5**, 1174 (1957).

[10] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[11] F. J. Owens and J. Charles P. Poole, *The new superconductors* (Plenum Press, 1996).

[12] J. Nagamatsu *et al.*, Nature **410**, 63 (2001).

[13] T. Ishiguro, K. Yamaji, and G. Saito, *Organic Superconductors* (Springer, 1998).

[14] C. Pfleiderer, Rev. Mod. Phys. **81**, 1551 (2009).

[15] A. Hebard, Nature **350**, 600 (1991).

[16] A. Y. Ganin, Nature Materials **7**, 367 (2008).

[17] http://en.wikipedia.org/wiki/Superconductivity .

[18] M. K. Wu *et al.*, Phys. Rev. Lett. **58**, 908 (1987).

[19] A. Schilling, M. Cantoni, J. D. Guo, and H. R. Ott, Nature **363**, 56 (1993).

[20] P. C. Canfield, Nature Materials **10**, 259 (2011).

[21] K. A. Moler, Nature **468**, 643 (2010).

[22] V. J. Emery, S. A. Kivelson, and J. M. Tranquada, Proc. Natl. Acad. Sci. USA **96**, 8814 (1999).

[23] D. C. Johnston, Advances in Physics **59**, 803 (2010).

[24] L. Fang *et al.*, Phys. Rev. B **80**, 140508(R) (2009).

[25] S. Nandi *et al.*, Phys. Rev. Lett. **104**, 057006 (2010).

[26] C. de la Cruz *et al.*, Nature **453**, 899 (2008).

[27] Y. Su *et al.*, Phys. Rev. B **79**, 064504 (2009).

[28] Y. Xiao *et al.*, Phys. Rev. B **80**, 174424 (2009).

[29] S. Nandi *et al.*, Phys. Rev. B **84**, 054419 (2011).

[30] J. Zhao *et al.*, Nature Physics **5**, 555 (2009).

[31] A. D. Christianson *et al.*, Nature **456**, 930 (2008).

[32] D. S. Inosov *et al.*, Nature Physics **6**, 178 (2010).

[33] M. Zbiri *et al.*, Phys. Rev. B **79**, 064511 (2009).

[34] R. Mittal *et al.*, Phys. Rev. Lett. **102**, 217001 (2009).

[35] L. Boeri *et al.*, Phys. Rev. Lett. **101**, 4 (2008).

[36] I. I. Mazin *et al.*, Phys. Rev. Lett. **101**, 057003 (2008).

[37] H. Ding *et al.*, Europhys. Lett. **83**, 47001 (2008).

[38] R. P. Huebener, N. Schopohl, and G. E. Volovik, *Vortices in unconventional superconductors and superfluids* (Springer, 2002).

[39] M. R. Eskildsen, Front. Phys. **6**, 398 (2011).

[40] J. S. White *et al.*, Phys. Rev. Lett. **102**, 097001 (2009).

[41] A. Furrer, *Neutron Scattering Investigations of Charge Inhomogeneities and the Pseudogap State in High-Temperature Superconductors* (Springer, 2005).

[42] S. Chi *et al.*, Phys. Rev. Lett. **101**, 217002 (2008).

E 2 Polymers: Structure

W. Pyckhout-Hintzen
Jülich Centre for Neutron Science 1 &
Institute for Complex Systems 1
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

A topical application of this spring school which deals with every day life is polymers. Polymers have a very rich behaviour and are ideal for investigations by means of neutron scattering techniques. As reported in the appropriate scattering introduction, this is due to the natural difference between a proton and a deuteron. Although chemically virtually no other properties are induced if some H's are replaced by some D's, this labelling allows to study e.g. the effect of different environments on the structure of polymers, on parts of the chains and so on. This is a strong advantage over scattering by xrays which is limited to systems which differ considerably in electron density. The H/D labelling does not change the xray scattering patterns. Nevertheless, the latter probe is very often used as a valuable complementary technique, especially in multiphase systems.

The following small treatise on polymers focuses only on the determination of their structure using the technique of Small Angle Neutron Scattering (SANS) and by no means is meant to be complete in the frame work of the Spring School. After a short introduction into general conformations of polymers which is independent of the monomer chemistry, thus generally applicable, we will summarize some structural models which describe the statistical properties of polymeric chains. Then we will develop some useful scattering approaches for mixtures of labelled and unlabelled chains, for linear and architecturally different chains. Details of the method of Small Angle Scattering are described at full length in the contribution by Frielinghaus (C1) and will not be repeated. Overall, this chapter tries to bridge the general text book science for linear polymers to more advanced and architecturally complex polymer structures in such a way that an appropriate base to facilitate the understanding of recent literature is provided . Therefore, the reader is made familiar with the access by scattering for the investigation of amorphous polymers, mostly in the melt state. A good introduction into general polymer physics is given in [1-7] and references therein. The connection to scattering methods is best treated in [8-11]. We warmly suggest these references for further lecture.

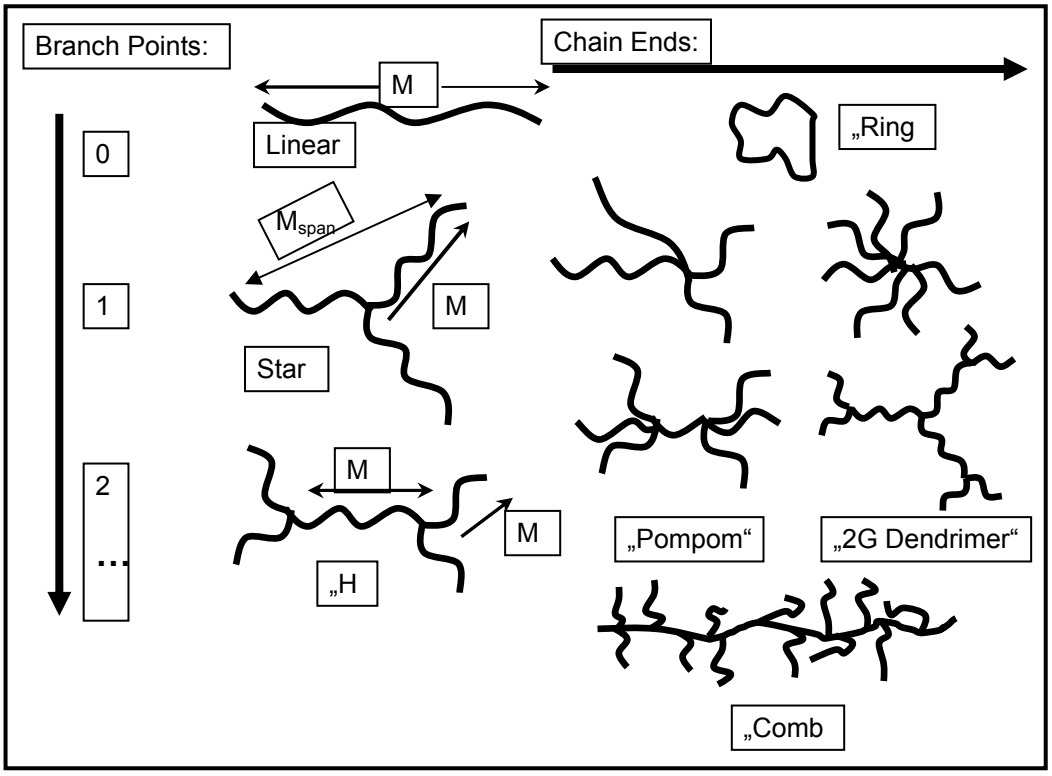


Fig. 1: *Overview of most important architectural polymer structures as a function of the number of branching points. Of these, the linear and the star polymer represent basic building blocks for higher complex branched polymers.*

2 Polymer Chain Models and Architecture

A polymer is a chain of several polyatomic units called monomers covalently bonded together. Since virtual all kinds of molecules can act as a monomeric unit, thereby only differing in the ways in which they can be bound together, a wealth of synthetic and naturally occurring polymers with enormous diversity in properties is nowadays known. We mention e.g proteins, DNA, glass, thermoplasts and rubber. They all belong to the class of polymers. Given their importance, an adequate description of model polymers has become a prerequisite. This brief summary serves to introduce the reader to some of the basic models with respect to scattering. The simplest descriptions of single-molecule models can then be expanded to branched polymers with some minor changes.

Polymers are often looked at like spaghetti-like or coiled molecules. Their stiffness varies from very flexible to rigid in the case of rods. Their configuration i.e their spatial distribution of segments changes all the time as the result of brownian dynamics which therefore leads to an enormous amount of possible configurations itself. The motion and dynamics of polymers will be discussed at full length in the session on polymer dynamics by Richter (E3). Here, we will stick to the determination and static investigation of their time-averaged structures.

To describe the statistical properties of such macromolecules, some variables need to be introduced. How can we calculate the size? For this, basic models were developed in the past

for linear chains with different levels of sophistication and approximations. We will summarize these only here in this chapter on the scattering of polymer chains in equilibrium and introduce at this stage also the corresponding length scales.

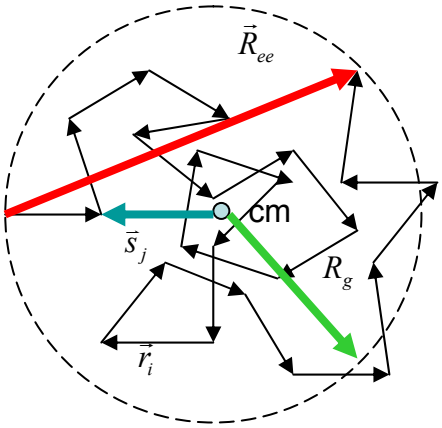


Fig. 2: A freely joined chain and definitions (see text).

The simplest model to deal with is the freely jointed chain (FJC). An example for such a configuration is shown in Fig. 2 and illustrates the idea that polymer chains are performing random walks. With this, the ensemble of systems can be characterized by average quantities. One is the so-called end-to-end vector.

$$\vec{R}_{ee} = \sum_{i=1}^{n} \vec{r}_i$$
$$R_{ee}^2 = \sum_{i,j=1}^{n} \vec{r}_i \vec{r}_j \qquad (1)$$

with R_{ee}^2 as the scalar quantity. The bond vector length (or step), connecting 2 atoms is l.
A further important parameter which has its counterpart in classical mechanics is the radius of gyration R_g :

$$R_g^2 = \frac{1}{n+1} \sum_{i=0}^{n} \vec{s}_i^{\,2} \qquad (2)$$

where now the vectors $\vec{s}$ are the distance vectors with respect to the center of mass, R_{CM} with

$$R_{CM} = 1/N \sum_{i=1}^{N} r_i \qquad (3)$$

Averaging over all conformations in both upper formulas leads for the case of a fully freely jointed chain to a mean squared end-to-end distance of

$$< R^2 >= \sum_{i,j=1}^{n} < \vec{r}_i \vec{r}_j >= nl^2 + 2\sum_{i<j}^{n} < \vec{r}_i \vec{r}_j > \quad (4)$$

whereas the average end-to-end distance $< R_{ee} >= 0$ due to the ensemble average. With the Lagrange theorem a relation between $< r_{ij}^2 >$ i.e. the squared distance between 2 scatterers *i* and *j* within the chain can be obtained. It states that

$$R_g^2 = \frac{1}{2(n+1)^2} \sum_{i,j}^{n} < (r_i - r_j)^2 >= \frac{1}{2(n+1)^2} \sum_{i,j}^{n} < r_{ij}^2 > \quad (5)$$

The calculation of the end-to-end distance which characterizes the full size of a chain i.e. it constitutes the diameter of a hypothetical sphere including the full chain, can now be performed for different chain models. The upper model of the FJC freely jointed chain assumes an equal probability in 3D of all bond vectors, random bond rotation angles while keeping the bond length constant. The orientation of each segment or bond is therefore independent of all others and as a consequence the 2nd term which is the scalar product in Eq. 4 averages out to zero. The mean radius of gyration can then be obtained from evaluating the double sum in Eq. 5 and using $r_{ij}^2 = (j-i)l^2$ - a result of random walk statistics -becomes

$$R_g^2 = \frac{1}{6} nl^2 \quad (6)$$

From both end-to-end distance and radius of gyration an important statement which applies to polymer chains in the bulk or theta state is already observed: the size depends on the square root of the number of basic steps, $\sqrt{n}$. This is a result which has its equivalent in the random walk statistics where now the position of the monomers is replaced by the trajectory of a randomly-diffusing particle and so the variable becomes the time, $\sqrt{t}$.

It is clear that the size or overall dimension of architecturally more complex branched chains cannot be estimated in the same way, basing on the end-to-end distances. This will involve a little bit different treatment as will be shown later.

However, real polymers are also not connected in this freely jointed way. Instead, bond angles assume well-defined values and correlations between bonds within the chain as well as from the environment will affect the size. Intuitively it can be understood that these correlations will vanish with increasing separation distance. Also, the flexibility of polymer chains is restricted by the fact that rotations with the so-called bond rotation angle φ_i are enabled within a certain range only due to steric reasons.

If we take the first FJC model as the reference then the difference to real polymers will be quantified in terms of a parameter which is denoted C_∞ . This is defined as

$$C_\infty = \frac{< R_{ee}^2 >}{nl^2} \quad (7)$$

In the case of the random walking chain in the FJC, C_∞ is thus 1 per definition.

A first refinement is the case of the freely rotating chain model (FRC) in which now valence angles are restricted additionally but still torsional angles are left free to rotate. Then the flexibility parameter after some calculations becomes

$$C_\infty = \frac{1+ < \cos\theta >}{1- < \cos\theta >} \tag{8}$$

which leads to values between 2 and 3 for the typical polymers. θ is the complementary angle between 2 bonds.

A further extension is the chain model with hindered rotation (HRC) which is better known as the rotational isomeric state model by Flory. This adds restrictions now also on the torsional degree of freedom. Instead, using the rotational barriers, a good estimate for allowed, energetically and sterically different configurations can be readily obtained. For C_∞, values considerably larger than 3 are found. Experimentally, the parameter as determined from small angle neutron scattering investigations is situated in the range between 3 and 10. We can summarize this by stating that highy-coiled polymers are characterized by low C_∞ values whereas extended or loosely-coiled chains show, on the contrary, high C_∞.

The description of chain dimensions thus involves several steps. A summary of currently used length scales in the literature which are inherently connected to the random coil description and resolution therefore is at its place here. Various expressions found for the chain end-to-end distance are

$$\begin{aligned} < R_{ee}^2 > &= C_\infty N_{mono} n_b l^2 \\ < R_{ee}^2 > &= N_{mono} l_{st}^2 \\ < R_{ee}^2 > &= N_{Kuhn} l_{Kuhn}^2 \end{aligned} \tag{9}$$

Here, n_b is the number of backbone bonds per monomer with rms bond length l, l_{st} is the statistical segment step length per monomer and l_{Kuhn} is the Kuhn segment step length. In the chain model of Kuhn [24], the correlation between bonds is taken into account so that the statistics of the FJC is retrieved for lengths larger than the Kuhn segment length. For this purpose, this effective segment length, l_{Kuhn} is introduced, which is related by $l_{Kuhn} = C_\infty l$. Likewise, $N_{Kuhn} = N / C_\infty$. The Kuhn segment is thus the minimum length scale above which a real chain behaves again like the freely jointed one. It further preserves the contour length with $l_{Kuhn} N_{Kuhn} = Nl$. Due to the direct comparison of molecular weights, provided by the synthetic chemists, with computations for the end-to-end distance, the definition of the statistical segment length per monomer is slightly more favoured whereas this is, however, no strict rule. This is also true for the following sections in this chapter on the scattering of polymers. Without going into detail and referring to more general literature [8], also a so-called persistence length l_p can be defined which is $2l_{Kuhn}$.

The distribution of the end-to-end vector is treated in many text books on statistical physics and will therefore not be repeated here. It shows that the distribution function *p(R)* of *any* intramolecular distance follows a Gaussian distribution with

$$p(R,N) = \left(\frac{3}{2\pi < R^2 >} \right)^{3/2} \exp\left(\frac{-3R^2}{2 < R^2 >} \right) \tag{10}$$

3 Scattering: intra- and inter-chain contributions

3.1 The single-chain structure factor

A scattering volume contains many polymer chains, each with *N* scatterers i.e. monomeric units here and with a coherent scattering length *b*. Their density is not constant and shows random fluctuations around an average value. $n_i(r) = 1$ if the monomer of type *i* sits at r_i. We define $< n(r) >= n = N/V$. The variation in the density is then $\Delta n(r) = n_i(r) - < n >$. The static structure factor is defined as the density-density correlation function $< n(-q)n(+q) >$. So, if in the system n_p chains, which all have *N* monomers with a scattering length *b*, are present the coherent scattering, i.e. the macroscopic differential cross section per unit volume, $\frac{d\Sigma}{d\Omega}(q) = \frac{1}{V_s}\frac{d\sigma}{d\Omega}(q)$ in [cm^{-1}] and further short *I(q)* can be calculated using:

$$I(q) = \frac{b^2}{V_s}\sum_{\alpha,\beta=1}^{n_p}\sum_{m.l=1}^{N} < \exp(i\vec{q}(\vec{r}_{\alpha,m} - \vec{r}_{\beta,l})) >= \frac{b^2}{V_s}S(q) \tag{11}$$

In the double summations the indices *m* and *l* are the monomer numbers and symbol α and β are different chains. $r_{\alpha,m}$ is thus the position of the *m*-th monomer on chain α. This can be rewritten into 2 parts: i.e. the intra-chain scattering is due to the contribution from 2 monomers on the same chain and an inter-chain part arises in the case that both monomers are situated on 2 different chains. For the simplest case of a dilute dispersion, we can further assume

$$< \exp(iq(r_{\alpha,m} - r_{\beta,l})) >= 0 \quad \Rightarrow \sum_{\alpha,\beta}\ldots \rightarrow \sum_{\alpha} \quad = n_p \tag{12}$$

because the chains are sufficiently well separated in space and the phase factor from each chain will be completely uncorrelated. Then the contribution of the inter-chain term cancels out on average. This allows us to determine the single chain structure factor $S_1(q)$. We can then re-write Eq. 11 (n_p and double sum over α,β cancel) as

$$S_1(q) = \frac{1}{N}\sum_{i,j=1}^{N} < \exp(i\vec{q}(\vec{r}_{m,i} - \vec{r}_{n,j})) > \tag{13}$$

We will come back to this in the context of the scattering of concentrated blends where inter-chain contributions have to be included in the full description. At this point it suffices to evaluate for pedagogical reasons the low *q* expansion. From this the radius of gyration which was introduced in the first part of this contribution, independent of the shape or structure can be determined. The Taylor expansion of the structure factor for $q << $ then yields

$$
\begin{aligned}
NS(q) &= \sum_{m,l} < 1 + i\vec{q}(\vec{r}_m - \vec{r}_l) - \frac{1}{2} q^2 \sum_{m,l} (r_m - r_l)^2 + ... > \\
&= N^2 - \frac{1}{2} q^2 \sum_{m,l} (< (x_m - x_l)^2 + (y_m - y_{\ l})^2 + ... > \\
&=> N^2 - \frac{1}{2} q^2 \sum_{m,l} \frac{1}{3} < (r_m - r_l)^2 > + ... \\
&= N^2 (1 - \frac{1}{3} q^2 R_g^2 + ...)
\end{aligned}
\tag{14}
$$

using the former result for R_g. Eq. 14 clearly proves that the low angle limit of the structure factor always yields both N and R_g if We will come back to this result later.

3.2 The Debye function

We have stated that the Gaussian distribution is a very good approximation for the real chain statistics. Then this yields

$$< \exp(i\vec{q}\vec{r}_{ij}) > = \exp(-1/2 < (qr_{ij})^2 > = \exp(-1/6 q^2 < r_{ij}^2 >) \tag{15}$$

For the isotropic gaussian chain we find

$$< (x_i - x_j)^2 > = < (y_i - y_j)^2 > = < (z_i - z_j)^2 > = \frac{1}{3} l_{st}^2 |i - j| \tag{16}$$

This result can be used and Eq. 13 can be evaluated in the full q-range with

$$S_1(q) = \frac{1}{N} \sum_{m,l}^{N} < \exp(i\vec{q}(\vec{r}_m - \vec{r}_l)) > = \frac{1}{N} \int_0^N dm \int_0^N dl \exp(-\frac{1}{6} q^2 l_{st}^{\ 2} (m - l)) \tag{17}$$

Here, the discrete sum has been already replaced by the continuous integral form (for N large enough, typically $N > 80$) which is the basis for all further calculations in this manuscript chapter on polymer scattering. It is the basic result for Gaussian chains. Its strength will be shown on selected examples, covering the most important fields of current polymer investigations. The statistical segment length per monomer l_{st} is used in conjunction with the total number of monomers, m and l, given by N. This integral can be solved analytically. The form factor $P(q)$ is defined as $S_1(q)/N$ and in discrete form is

$$P(q) = \frac{1}{N^2} \sum_{m,l} \exp(-\frac{1}{6} q^2 l_{st}^2 (m - l)) \tag{18}$$

For long chains ($N \rightarrow \infty$), $P(q)$ is called the Debye function $g_D(x)$ with

$$g_D(x) = \frac{2}{x^2} (\exp(-x) - 1 + x) \tag{19}$$

and the argument $x = (qR_g)^2$. A SANS experiment with fit to the Debye function is presented in Fig. 3 in a linear scale and in the form of a 2nd moment Kratky representation achieved when $g_D(x)$ is multiplied by q^2. The Kratky representation emphasises the high q-regime. For Gaussian chains with an asymptotic q^{-2} behaviour the high q regime then assumes a plateau. As shown the data are in perfect agreement with the Gaussian chain results derived above.

Small and high q regimes can also be explicitly obtained from expanding Eq. 19 for $x \to 0$. For small q, we then have $g_D(x) \approx 1 - 1/3x + ... = 1 - 1/3q^2R_g^2 + ...$. Reversely, this is the Taylor equivalent of *exp(-x),* which can be conveniently plotted in either the Guinier representation *ln(g_D) vs* q^2 from which then the slope contains R_g, or in a Zimm plot i.e. the inverse structure factor *1/S(q) vs.* q^2. The disadvantage of the latter is that slope and intercept in the linearized form are coupled whereas they are independent in the logarithmic way. The low q expansion allows a fast determination of R_g. An ultra-fast single-point estimate of the chain dimension can be obtained as follows: for $q^* = 1/R_g$ it can be easily obtained then that the scattering intensity at q^* has dropped to *2/3* of the forward scattering value at *q=0*.
Thus, in the low q regime we obtain direct information on the chain molecular weight M_w (via N) and the radius of gyration R_g. For high q, Eq. 19 on the other hand gives

$$g_D(x) \approx \frac{2}{x} = \frac{2}{q^2R_g^2} \sim \frac{1}{C_\infty} \qquad (20)$$

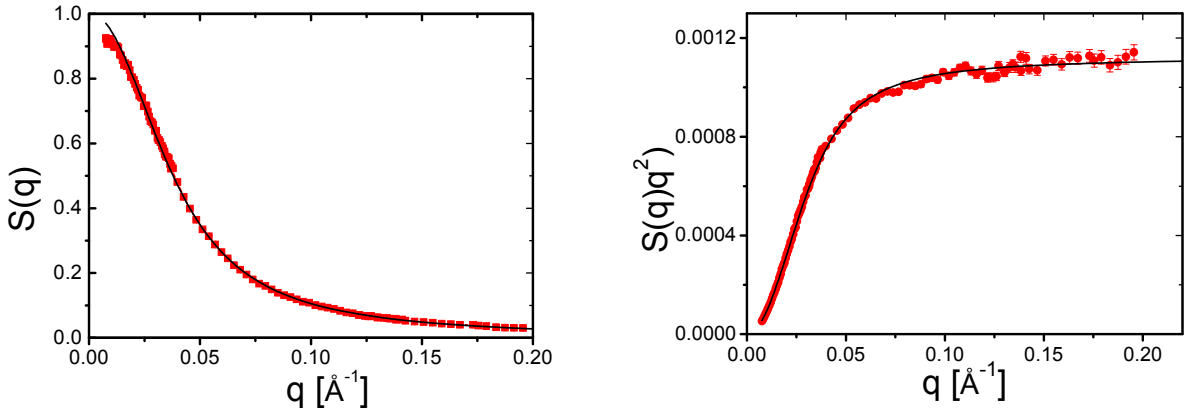


Fig. 3: *Experimental scattering data, normalized to S(q) on a poly-alkylene-oxide polymer, obtained at KWS2@FRM-2, Munich [15]. On the left, the Debye curve fits ideally, showing the Guinier region and the high q-dependence* q^{-2}. *On the right, the Kratky region is highlighted.*

However, since data at high q, where the level of incoherent background (see D1) plays an important role, can be sensitively affected by erroneous subtraction, its use for the estimation of R_g is limited.

4 Blend of linear Polymers

4.1 Mixture of 2 polymers

The above applied to a general system in which the contrast was given by a polymer vs. a background with zero scattering length. This is not what one wants to measure nor is it generally the case. The power of the neutron scattering approach is the use of the natural contrast between protonated and deuterated components, i.e. where a label is able to provide

the structure of a polymer chain in the phase which scattered before incoherently as from a single phase. The total structure factor in a dense system is obtained from Eq. 11 and split into an intra-chain and an inter-chain part.

$$S(q) = MN^2P(q) + M(M-1)N^2R(q) \tag{21}$$

If we consider a mixture of 2 polymers with M chains, which are identical in length with N monomers and have no preferential interactions with each other, we define the form factor $P(q)$ and the inter-chain structure factor $R(q)$ which is sometimes even playing a dominating role. $P(q)$ has been defined before and R(q) is formulated similarly with double sums:

$$\begin{aligned} P(q) &= \frac{1}{N^2}\sum_{m,l}^{N} < \exp(iq(r_{1,m} - r_{1,l})) > \\ R(q) &= \frac{1}{N^2}\sum_{m,l}^{N} < \exp(iq(r_{1,m} - r_{2,l})) > \end{aligned} \tag{22}$$

The total structure factor then becomes (for $N_H = N_D$):

$$I(q) = \frac{1}{V_s}\sum_{i}^{2}\sum_{j}^{2} b_i b_j S_{ij}(q) \tag{23}$$

With for $i=H$ and $j=D$

$$\begin{aligned} S_{DD}(q) &= M_D N_D^2 P_D(q) + M_D^2 N_D^2 R_D(q) \\ S_{HH}(q) &= M_H N_H^2 P_H(q) + M_H^2 N_H^2 R_H(q) \\ S_{HD}(q) &= M_D M_H N_H N_D R_{HD}(q) \end{aligned} \tag{24}$$

With the incompressibility hypothesis one gets to

$$\begin{aligned} S_{DD} + S_{DH} = 0 \rightarrow S_{DD} &= -S_{HD} \\ S_{DD} &= -S_{HH} \end{aligned} \tag{25}$$

We have implicitly assumed that the hydrogenous and deuterated polymers are identical and therefore exhibit the same intra-chain $P(q)$ and inter-chain structure factor $R(q)$, independent of the isotopic labelling. The inter-chain contribution can be expressed in terms of the intra-chain contributions which simplifies the scattering intensity to a rather simple expression in terms of a single chain contribution factor only.

$$\begin{aligned} I(q) &= \frac{1}{V_s}(b_D^2 S_{DD}(q) + b_H^2 S_{HH}(q) + 2b_D b_H S_{HD}(q)) \\ I(q) &= \frac{(b_D - b_H)^2}{V_s} S_{DD}(q) \end{aligned} \tag{26}$$

4.2 Mixture of 2 polymers in a matrix or solvent

The same reasoning as in the former blend can be done if it is mixed with a third component which may be e.g a small scale solvent, a macromolecule, particle or a complex environment. Latter can consist of e.g. a hard confinement or aggregated nanoparticles or soft network.

Working through the equations one gets to a similar expression as for the simple symmetric blend with the exception that additional freedom or contrast can be provided. The contrast situation has been treated in detail in the scattering section (C1) and we refer to that. Taking into account that the medium in which the blend is embedded in, has a non-zero scattering length b_0, the blend can be mixed as to obey

$$b_0 = xb_D + (1-x)b_H \tag{27}$$

provided b_0 takes a value between both constituent polymers with volume fractions x and $1\text{-}x$. It can itself be a mixture of e.g. different solvent molecules that do not show a coherent structure on the length scale of the macromolecules itself or e.g. also a random co-polymer with matched scattering length length. The full coherent intensity becomes

$$\begin{aligned} I(q) &= \frac{1}{V_s}((b_D - b_0)^2 S_{DD}(q) + (b_H - b_0)^2 S_{HH}(q) + 2(b_D - b_0)(b_H - b_0)S_{HD}(q) \\ &= \frac{1}{V_s}[(b_D - b_H)^2 x(1-x)M_D N_D{}^2 P_D(q) + (xb_D - (1-x)b_H - b_0)^2 S_{TOT}(q) \end{aligned} \tag{28}$$

where we have substituted the appropriate form factors already. Eq. 28 shows different possibilities which can be advantageous and which are applied in several approaches, at least qualitatively. We notice that the single chain form factor of the blend can still be measured if

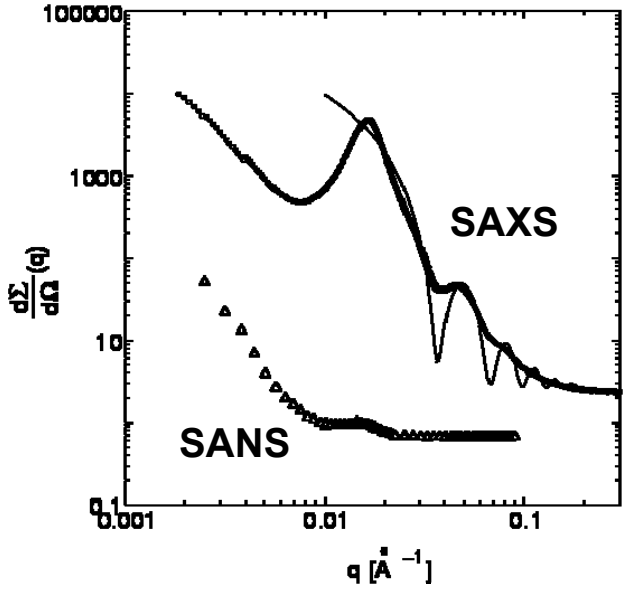


Fig. 4: *Experimental phase matching: the scattering for neutrons is almost extinguished whereas xrays show the phase structure, i.e a cylindrical polystyrene mesophase domain inside a sea of polyisoprene arms [unpublished, PhD Thesis A. Botti, Univ. of Münster,2001].*

the 2nd term is known or is matched out. The contrast matching can be achieved if the blend is compositionally mixed so that the contrast factor is zero. The scattering length of the 3rd component or medium b_0 itself can be an effective one if e.g. solvent molecules are used which can be adjusted as to cope with the composition. Eq. 28 has another consequence: if

both polymers have the same scattering length, the 1st term vanishes and the intensity is given by the contrast with the medium. The 2nd term simulates a system in which all components have the same scattering length <*b*> and thus concentration fluctuations of the polymer become visible. This is the situation which is present if the probe is not neutrons but xrays. Likewise, the first term also disappears if the mixture consists of only 1 single type of polymer! Latter is the direct equivalent of the ray study. Now the phase in which the polymer is can be studied. S_{TOT} corresponds then to e.g. the nanocomposite structure or the network-like environment. An example is given in Fig. 4. We note that, still using the same incompressibility rules, Eq. 28 can be re-casted such that then the 3 different diagonal structure factors, i.e S_{HH}, S_{DD} and S_{TOT} occur.

4.3 The dilute solution

To be complete we mention only the case of dilute polymer solutions. For low concentrations of labelled species in one of the components, Zimm has proposed an effective single contact model. If 2 monomers on closeby chains interact, their coordinates can be distributed in the interchain structure factor like

$$\sum_{ij}^{N} < \exp(iqr_{1i2j}) > = \sum_{ij}^{N} < \exp(iqr_{1i1k}) > < \exp(iqr_{1k2m}) >< \exp(iqr_{2m2j}) > \tag{29}$$

The middle term is the excluded volume between both polymer chains.. If this is substituted in *R(q)* and terms are reorganized, Zimm derived that

$$R(q) = -2A_2M^2P^2(q) \tag{30}$$

A2 is the 2nd virial coefficient. The famous Zimm equation follows :

$$\frac{Kc}{I(q)} = \frac{1}{M_wP(q)} + 2A_2c + ... \tag{31}$$

We will not go into details in the use of this equation which was significantly used in the last decades for the determination of polymer chain structure and polymer-solvent interactions and we refer to some specialized literature. For blends its use has been marginal and for these the high concentration method is more favoured. It can be seen from the functional form in Eq. 31 that the single chain information can be extracted from extrapolation to *c=0* whereas the 2nd virial coefficient can be derived involving the extrapolation of *P(q)* to *P(q=0)* where this is by definition =1.

4.4 Mixture of 2 polymers with interaction

In the former 2-component mixture, polymers behaved ideal, did not interact much and especially were symmetric. For asymmetric systems, one obtains the result in Eq. 32 which is identical to that obtained from a complicated random phase approximation (RPA) by Leibler [12] which applies to a general mixture of 2 asymmetrical polymers that differ only in their isotopic labelling and possibly interact. In the classical approach the Flory-Huggins interaction parameter is therefore set to $\chi = 0$. We get

$$S(q) = \frac{M_H N_H^2 P_H(q) M_D N_D^2 P_D(q)}{M_H N_N^2 P_H(q) + M_D N_D^2 P_D(q)} \Rightarrow [\frac{1}{S_{0HH}(q)} + \frac{1}{S_{0DD}(q)} - 2\chi_{HD}]^{-1} \quad (32)$$

The index *0* now means that we deal only with the bare correlation functions assuming no interaction between the chains and is not to be mixed up with *S(q)*. S_{0HH} (and S_{0DD} similarly) is defined as $\phi N_H g_D(N_H)$. ϕ is the respective volume fraction in the mixture. The χ parameter is strongly temperature dependent. For a mixture of 2 polymers, by incompressibility the cross term S_{0HD} is zero. In the random phase approximation the chains are nearly ideal and perform random walks. The full RPA treatment (Eq. 33) and the consequences becomes clear if the mixture is replaced by e.g. a tri-block copolymer. It can be shown by exercise that the simple mixture can be retrieved by substituting the appropriate function into the partial structure factors. The RPA approximation thus effectively accounts for the role of the inter-chain contributions.

5 Tri-block copolymers: an example for RPA

The random phase approximation (RPA) can be used for interacting systems and was originally developed by P.G. De Gennes. If different chemistries come together, also different interactions can be expected and therefore phase separating mechanisms can get active. Even the exchange between H and D can be sufficient. Since scattering is just sensitive to fluctuations in the composition, it is therefore very often used to study phase diagrams. Let us consider here the scattering of a symmetrical tri-block copolymer structure, consisting of identical monomers but differing only in their scattering length. We assume ideal mixing, thus negligible contributions from the Flory-Huggins parameter χ. We refer to the literature for more general treatments in isotopic block copolymers or even general block copolymers of different chemistries which then lead to micro-phase separation and meso-phases. They are discussed in the section on the SANS technique. The present example treats a HDH tri-block with a total of N monomers with N_H resp N_D monomers in the separate blocks of which the central block is deuterated. Then the problem reduces to the calculation of the partial structure factors which is done similar as we did to obtain the Debye function g_D.
These are fed into the general RPA equation, Eq. 33 which can re-organized to highlight the different contributions

$$\begin{aligned} S_{RPA}(q) &= \frac{S_{0DD}(q)S_{0HH}(q) - S_{0HD}^2(q)}{S_{0DD}(q) + S_{0HH}(q) + 2S_{0HD}(q)} \\ &= S_{0DD}(q) - \frac{(S_{0DD}(q) + S_{0DH}(q))^2}{S_{0DD}(q) + S_{0HH}(q) + 2S_{0HD}(q)} \end{aligned} \quad (33)$$

We here explicitly assumed no interactions and near-ideal Gaussian chains. *f* is the fraction of the *i*th monomer type in the tri-block, $f=N_i/N$ and N_i is the number of segments in the block.
Performing the calculations leads to new correlations which are absent for a mixture of simple linear polymers by adjusting the limits of the integration:

$$S_{0H\ldots H}(q) \rightarrow \int_{0}^{N_H} \int_{N_H+N_D}^{N} = \left(\frac{1-\exp(-q^2R_{gH}^2)}{q^2R_{gH}^2}\right)^2 \exp(-q^2R_{gD}^2)$$
$$S_{0HD}(q) \rightarrow \int_{0}^{N_H} \int_{N_H}^{N_H+N_D} = \left(\frac{1-\exp(-q^2R_{gH}^2)}{q^2R_{gH}^2}\right)\left(\frac{1-\exp(-q^2R_{gD}^2)}{q^2R_{gD}^2}\right) \quad (34)$$

Here, R_{gD} resp R_{gH} are defined as $\sqrt{f}\,R_g$ resp $\sqrt{(1-f)/2}\;R_g$ in view of the gaussian character (remember the N-dependence of the end-to-end distance!). It can be easily seen that the RPA for the tri-block can be simplified back to the limiting case of a linear chain without labeled wings by letting the wings $\rightarrow 0$. Also, the RPA of the tri-block can be easily converted and transformed to the case of a general di-block copolymer by dropping the non-bonded *H…H* correlation term.

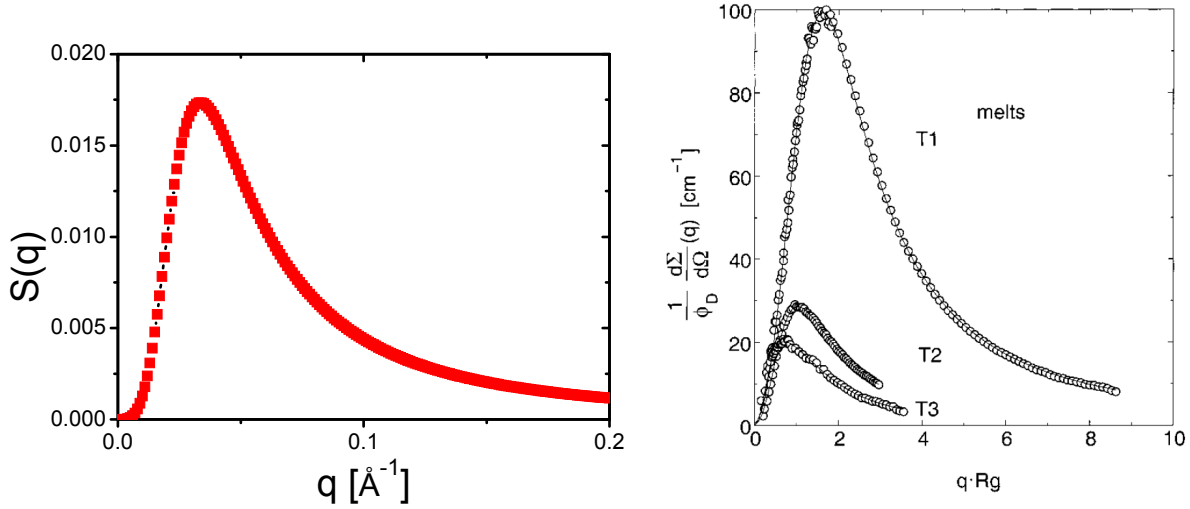


Fig. 5*: Computational example for a tri-block scattering with Rg~90Å (left) and on the right an experimental system obtained at KWS1@FRJ2, Jülich with fits to 3 different tri-blocks differing only in the middle block length in reduced representation using Rg of the center block [13].* Reprinted with permission from Macromolecules, ***34,*** 2186(2001). Copyright (2001) American Chemical Society.

Eq. 33 nicely illustrates that the scattering is dominated by 1 block. As an example we choose the D block. For an isolated polymer with the length N_D one expects a Debye-like curve. However, the 2nd term in eq 33 contains the correction to this due to the connectivity of this block now into a tri-block. The connectivity leads to the subtraction of the 2nd term and leads to a pronounced peak in the scattering intensity. The peak intensity and position is determined by the center block and the interaction parameter χ (if any) leading to stronger peaks close to (micro-)phase separation. This is the correlation hole effect. At *q=0* there are no composition fluctuations as on this length scale the same fluctuations as inside the molecules occur. For intermediate *q* the largest fluctuations occur around $q \sim 1/R_g$ of a branch and for even larger *q* there is no difference between a blend of unconnected blocks and the block copolymer. The

dependence is again that of the random walk. If this tri-block is mixed with a linear chain which has the scattering length in common with one of the blocks, the respective contributions are weighed with the volume fraction in the blend and added to the respective partial structure factors. This leads to new composition fluctuations at low q. In the original work by Leibler it is even shown that due to sometimes just polydispersity in the block copolymer, be it in the total length, or in the block length or composition, also a non-zero intensity is expected. Using the special anionic polymerization technique which guarantees very narrow molecular weight distributions and low polydispersities well below 1.1, this can be very well avoided as shown in the example from our lab (Fig. 5) and the analysis is a strong measure of quality of the synthetic efforts.

6 Branched Chains

6.1 The branching parameter g

There are many different architectures in branched chains. A summary of the most prominent ones which can be synthesized by the mentioned method of anionic polymerization as model systems are shown in Fig. 1. There, e.g a star-branched chain consists of a core and arms of a more or less similar length. Therewith, a two-arm star polymer is essentially a linear polymer. A comb polymer, on the other hand, consists of a linear-chain backbone and many combs or teeth that stick out from functional groups distributed somehow along the backbone. These junctions or main-chain branching points may be uniformly or randomly spaced. The comb length can be highly uniform depending on the synthesis route. Comb polymers with uniform distribution of branching points and a uniform distribution of branching length can be made with special techniques and low yields only. Branched chain**s** where branching points are distributed randomly, are much more common, however. Branches may have also further branches, leading to higher generations. Then the longest chain i.e. the backbone in the random-branched chain with the span molecular weight, can hardly be distinguished anymore. The length of the branches divides these systems then in long-chain branched and short-chain branched polymers. Long-chain branching reactions can be induced already in special chain-reaction polymerization methods. If the frequency of branching is high, then the chain becomes hyper-branched. This is another class. E.g in a dendrimer shown in Fig. 1, every repeating unit is trifunctional. Starting at the center, the number of segments or branches in a layer or generation increases by a factor of 2 in the next generation.

The best quantity to characterize the overall dimension of a branched chain is through the radius of gyration, R_g. An end-to-end distance cannot be uniquely identified anymore for non-linear chains and instead, a branching parameter g is defined as

$$g = \frac{R_{gb}^2}{R_{gl}^2} \tag{35}$$

where R_{gb}^2 now is the mean squared radius of gyration for the branched chain, and R_{gl}^2 is the mean square radius of gyration for the linear chain. The ratio is calculated for the two polymers of the same molecular weight. It can be easily seen in Fig. 1 that branching concentrates monomers around the center of mass. Therefore, $g \leq 1$ always. As an example we calculate here g for an n_A-arm star polymer. With some more efforts and approximations

the branching factor of more complex architectures can be obtained similarly. We assume that each arm behaves as an ideal chain and has the same number of monomers, N_1 (N_1 »). Let $\vec{r}_{ij}$ be the position of the jth monomer (j :0…N_1) on the ith arm (i:1…n_A), as illustrated in Fig. 6. With Eq. 5, the mean-squared radius of gyration R_{gb}^2 is expressed as

$$\begin{aligned} R_{gb}^2 &= \frac{1}{2n_a^2N^2}\sum_{i,n=1}^{n_A}\sum_{j,m=1}^{n_A}<(r_{i,j}-r_{m,n})^2> \\ &= \frac{1}{2n_a^2N^2}\sum_{i=1}^{n_A}\sum_{j,n=1}^{N}<(r_{ij}-r_{in})^2> + \sum_{i\neq m}^{n_A}\sum_{j,n=1}^{N}<(r_{ij}-r_{mn})^2> \\ &= \frac{1}{3}b^2n_AN(N+1)(3n_AN-2N-1) \end{aligned} \tag{36}$$

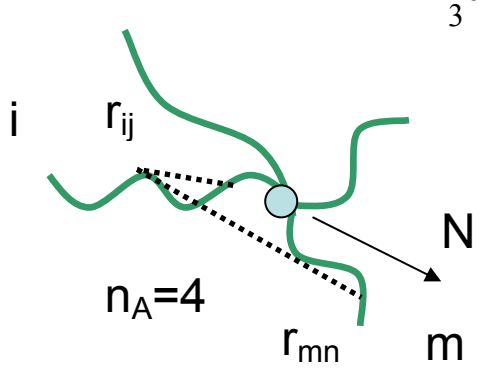


Fig. 6: *Definitions of coordinates for a symmetric 4-armed star polymer*

The mean squared monomer distance is calculated for monomers on the same arm and those on different arms separately. For large enough N we obtain

$$R_{gb}^2 \cong b^2\frac{N}{n_A}\left(\frac{1}{2}-\frac{1}{3n_A}\right) \tag{37}$$

Again, to prove the consistency the limiting case of the linear polymer is well suited. Linear polymers can be looked at as 2-arm stars or even as 1-arm star objects. As one can easily verify, the known relation for R_g coming from the FJC model is obtained. Thus, with this the branching parameter g is a nice descriptor for branching.

$$\begin{aligned} n_A = \{1,2\} \rightarrow R_{gl}^2 &= \frac{Nb^2}{6} \\ g = \frac{1}{n_A}(3-\frac{2}{n_A}) \quad &< 1 \end{aligned} \tag{38}$$

6.2 Form factor of a symmetric star polymer

We can now calculate the form factor $P_{\text{star}}(q)$ for this n_A-arm star polymer with this uniform arm length N_1. When calculating the average of $\exp(i\vec{q}(\vec{r}_m-\vec{r}_n))$, it is again necessary to distinguish between the two cases for r_m and r_n i.e. both on the same arm and both being on

different arms. The former case takes place with a probability of $1/n_A$ and obviously $(1-1/n_A)$ for the other. Then,

$$P_{star}(q) = \frac{1}{n_A} < \exp(iq(r_m - r_n)) >_1 + (1 - \frac{1}{n_A}) < \exp(iq(r_m - r_n)) >_2 \qquad (39)$$

where the subscripts 1 and 2 correspond to the two cases. The averages are taken with respect

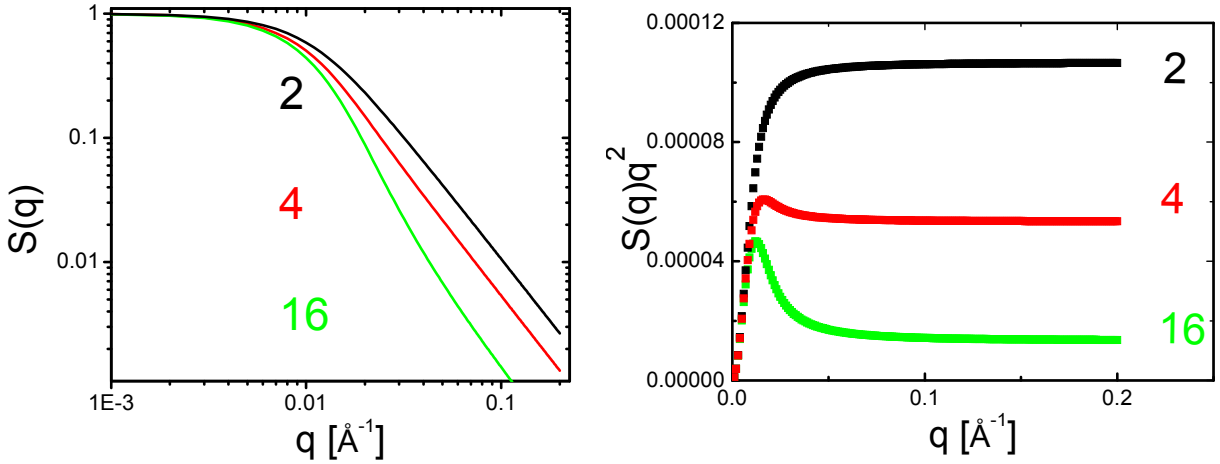


Fig. 7*: Star scattering: log-log and Kratky representation for stars with same number of monomers but varying functionality, going from 2 to 16. The results show clearly that for n_A=16 the scattering starts to resemble that of spherical objects from the apparent q^{-4} dependence below q=0.02Å^{-1} .*

to the two monomers over the length of the arm. The second average can be approximated after some manipulation to consist of the independent averages of single arms as

$$\begin{aligned} &< \exp(iq(r_m - r_n)) >_2 \; -\!\!- >< \exp(iq(r_m - r_0)) >^2 \\ &< \ldots >_2 = [\frac{1}{N_l} \int_0^{N_l} dn \exp(-\frac{1}{6} q^2 n l_{st}^{\,2})]^2 = [\frac{1 - \exp(-q^2 R_{g,l}^2)}{q^2 R_{g,l}^2}]^2 \end{aligned} \qquad (40)$$

Combining all results leads to the form factor of a star polymer with n_A arms and radius of gyration per arm $R_{g,l}$ with N_l monomers

$$P_{star}(q) = \frac{1}{n_a^2 N}(n_a N^2 g_D(N) + n_a(n_a - 1)[(\frac{1 - \exp(-q^2 R_g^2)}{q^2 R_g^2})]^2) \qquad (41)$$

6.3 Form factor of a ring polymer

Ring polymers are very interesting macromolecules in that they differ structurally and dynamically from the typical linear polymers due to the total absence of chain ends that are dominating for the dynamics in the melt. They have become of renewed interest in the last years due to the synthetic efforts in producing linear contaminant-free ring structures while

minimizing concatenation reactions. As Fig. 8 shows, there are 2 different ways to define the intra-chain correlation between two monomers *i* and *j*. Both ways are equally probable and the occurrence is taken into account by multiplying the probabilities of selecting each different path. For the mean-squared distance between 2 monomers and the corresponding form factor we obtain then in analogy with the linear chain and including the closure relation

$$\begin{aligned} &< r_{ij}^2 >= b^2 |i-j| (1 - \frac{|i-j|}{n}) \\ &\rightarrow P(q) = \frac{1}{n^2} n \sum_{i,j} \exp(-\frac{q^2 b^2 |i-j|}{6} (1 - \frac{|i-j|}{n})) \end{aligned} \tag{42}$$

The single chain structure factor differs in a sensitive way from the pure linear curve as can be seen in Fig. 8 . It leads to a peaked structure in the Kratky representation. This expresses the similarity of a ring polymer with a star or branched polymer (Fig. 7). The monomers of a ring are on average located at closer distance to the center of mass than in a linear chain, so an enhancement of the compacticity of the structure can be expected. The peak arises due to the increased correlation through the closed cycle and is therefore related to R_g of the ring. Latter can be calculated to be smaller than the linear by $\sqrt{2}$.

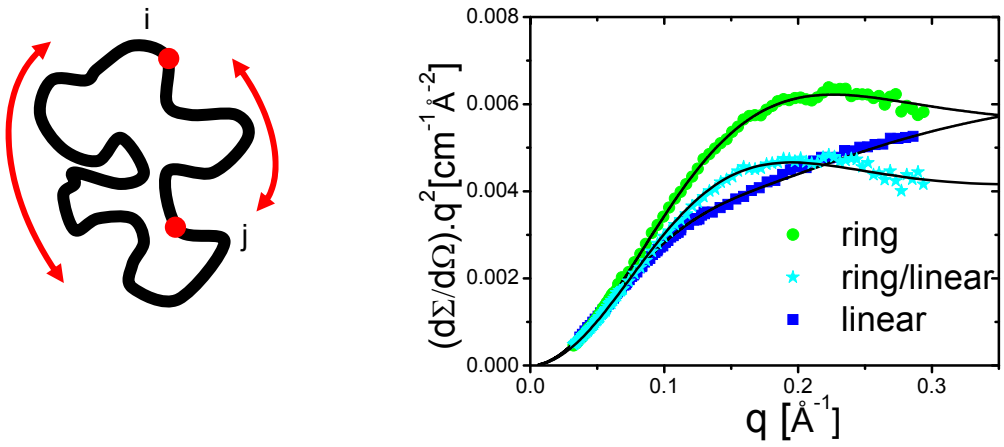


Fig. 8: *Ring scattering: 2 possible ways to define the distance r_{ij}. Kratky representation to highlight the difference in the structure due to ring closure. In the ring/linear mixture the peak shifts to lower q and indicates a swelling of the ring conformation by penetrating linear chains. Adapted from [14].*

7 Summary

We have discussed in a nut-shell various possibilities to investigate the structure of model polymeric chains in various environments and constitutions, some in more detail than others while keeping the focus on dense, amorphous and well-mixed systems. Mixtures of chemically different polymers behave much more complicated, are partly treated in other lectures and can be very well studied as well by the small angle xray scattering method. We have tried to summarize here only the most important concepts of the statistical structure of

ideally monodisperse chains and we have demonstrated how the particular structure affects the scattering behaviour. This text should be a basis in order to be of help in calculating more complex structures which are built up from the here discussed basic entities i.e the linear and the star polymer. It was a general aim thus to provide some useful basics which can be transferred to related sample systems.

Appendix

We had defined the macroscopic differential cross section $\frac{d\Sigma}{d\Omega}(q)$ per unit volume in [cm^{-1}] as $\frac{1}{V_s}\frac{d\sigma}{d\Omega}(q)$. Here, Vs is the irradiated sample volume. In this context, another very often used parameter, i.e. the scattering length density (or SLD), ρ, can be defined which replaces the scattering length *b*. This SLD or ρ is defined as $\sum_i b_i / V_0$ with V_0 a reference volume over which the scattering lengths *b* are summed up and normalized to its volume. For polymers, V_0 typically is the monomeric volume and for simple molecules or particles it can be the sum of all the constituents. With this the intensity becomes

$$I(q) = \frac{N_p}{V_s} N^2 V_0^2 \rho^2 P(q) = \frac{N_p}{V_s} V_p^2 \rho^2 P(q) \tag{43}$$

With the relationship $N_p / V_s = \phi / V_p$ and V_p and ϕ being the polymer or particle molecular volume and volume fraction, this can be further simplified. An expression for SANS intensities which is often encountered, especially among experimentalists is then found as

$$I(q) = \phi V_p \Delta\rho^2 P(q) \tag{44}$$

Here, $\Delta\rho^2$ represents the contrast factor i.e. the squared difference between the scattering length densities of both components.

References

[1]G. Strobl, *The Physics of Polymers* (Springer,1996)
[2]P.J. Flory, *Principles of Polymer Chemistry* (Cornell University Press,1953)
[3]P.G. DeGennes, *Scaling Concepts in Polymer Physics* (Cornell University Press,1979)
[4]M. Rubinstein, R. Colby, *Polymer Physics* (Oxford University Press,2003)
[5]W.W. Graessley, *Polymeric Liquids & Networks: Structure and Properties* (Taylor & Francis,2004)
[6]I. Teraoka, *Polymer Physics: an Introduction to Physical Properties* (John Wiley& Sons,2002)
[7]M. Doi, S.F. Edwards, *The theory of Polymer Dynamics* (Clarendon Press,1999)
[8]J.S. Higgins, H. Benoit, *Polymers and Neutron Scattering* (Clarendon Press,1994)
[9]B. Hammouda, http://www.ncnr.nist.gov/staff/hammouda/the_SANS_toolbox.pdf
[10] G.D. Wignall, *Neutron Scattering for Polymers*, in *Encyclopedia of Polymer Science and Engineering*, Vol 10 (John Wiley & Sons,1987)
[11] *J. Oberdisse, W. Pyckhout-Hintzen, E. Straube, Structure Determination of Polymer Nanocomposites by Small Angle Scattering*, in *Recent Advances in Polymer Nanocomposites (VSP,2009)*
[12] L. Leibler, H. Benoit*, Polymer, **22***, 195(1981)
[13] S. Westermann, W. Pyckhout-Hintzen, D. Richter, E, Straube, S. Egelhaaf, R. May, *Macromolecules*, **34**, 2186(2001)
[14] A. Bras, R. Pasquino, T. Koukoulas,G. Tsolou, O. Holderer, A. Radulescu, J. Allgaier, V. Mavrantzas, W. Pyckhout-Hintzen, A. Wischnewski, D. Vlassopoulos, D. Richter, *Soft Matter*, **7,** 11169(2011)
[15] C. Gerstl, G. J, Schneider, W. Pyckhout-Hintzen, J. Allgaier, S. Willbold, D. Hoffmann, U. Disko, H. Frielinghaus, D. Richter, *Macromolecules*, **44**, 6077(2011)

E 3 Polymerdynamics

D. Richter
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Look around the house and you will see that you are surrounded by many kinds of polymers: plastic containers, surface coatings in the kitchen, toys and clothes in the living room and bedroom. Modern equipment and components in your car and work place are often made of or coated with polymer composite materials. Today traditional materials, such as metals, ceramics and wood have partly been replaced by synthetic polymers which may be stronger, lighter and cheaper and which through scientific research can be tailored to specific requirements.

It is hard to imagine life without modern synthetic plastics and rubbers. These polymers can be moulded into almost any shape, extruded into thin films and fibres, applied as coatings and given bright colours or made transparent. New polymer composites are continually being developed including reinforced rubbers or construction materials even for aeroplanes.

Polymers are one of the most important products of chemical industry. The development of this industry in Germany started in the second half of the 19th century. BASF was founded in 1866. In 1885 it had already 2335 employees. In 1900 BASF had grown to 6771 employees. Similarly Bayer was founded in 1881. In 1885 Bayer employed 24 chemists and 300 workers. Just 11 years later it had grown to 104 chemists and 2644 workers. The 2006 turnout of chemicals in Germany is given in Table 1. Among these products polymers are on the third rank with 25.88 billion Euros. Thus, polymers are indeed a very important commodity.

Table 1:
Turnover along products (Germany 2006)

Product	Turn over [10^9 €]
Anorganic base chemicals	9,88
Petrochemicals	21,88
Polymers	25,88
Fine + special chemicals	32,62
Pharmaceuticals	29,33
Body care products	10,02

Other than synthetic polymers, biopolymers are highly complex molecular structures with a sharp molecular weight and a uniquely defined sequence of building blocks. They are the molecular manifestations of life with a high variety of structure and specific function. Nevertheless, at least in their denaturated state they exhibit similarities with synthetic materials while in their folded state the specific function evolves. Also the general principles of selfassembly in structure formation are common to both species.
The diffusional motions of long flexible polymers constitute fascinating physics and at the same time represent one of the great challenges of modern material science. The drive towards the molecular understanding of the complex viscoelastic properties of polymer liquids is the focal point of rheology and connects the classical chemical engineering approach with modern physics [1]. There the tube model invented by Doi and Edwards [2]

and de Gennes [3] has shown itself as the most successful molecular model describing the topological confinement imposed by the mutually interpenetrating polymer chains in the melt. In terms of this so called reptation model a theory of viscoelasticity has been developed that describes the main features of polymer melt rheology.

Large scale protein motions on the other hand are critical for proteins to coordinate precise biological function. Such dynamics are invoked in regulatory proteins, motor proteins, signalling proteins and structural proteins. Structural studies have documented the conformational flexibility in proteins accompanying their activity. Results from macroscopic studies such as biochemical kinetics and single molecule detection have also shown the importance of conformational dynamics and Brownian thermal fluctuations within the proteins or protein complexes. However, the time dependent dynamic processes that facilitate such protein motions remain poorly understood and experimentally nearly untouched [4].

This lecture aims to identify general principles of chain motion on a molecular scale which underpin the macroscopic properties and presents concepts and experimental results on these motional mechanisms in space and time. We will mainly address the dynamics of linear homopolymers and give one example for studies of the molecular dynamics of biopolymers.

2 Macroscopic Dynamics

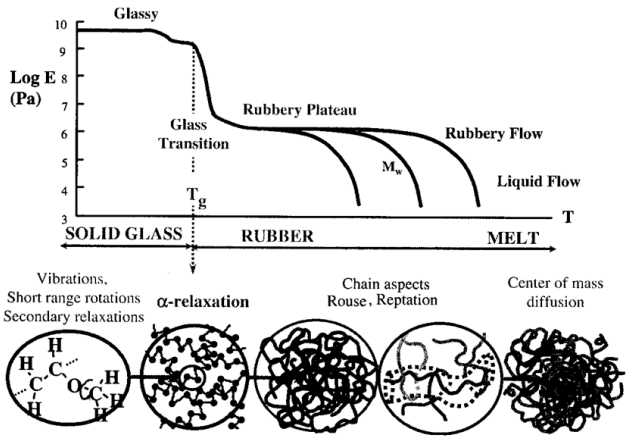


Fig. 1: *Schematic presentation of the modulus of a polymer melt as a function of temperature. The sketches below symbolise the types of motions that give rise to the macroscopic behaviour.*

Dynamic processes in polymers occur over a wide range of length and time scales. Fig. 1 relates the dynamic modulus, as it may be observed on a polymer melt, with the length and time scales of molecular motion underlying the rheological behaviour. Our example deals with an amorphous polymer excluding any crystallization processes. It is clear, that we can distinguish several different regimes. At low temperatures the material is in a glassy state and only small amplitude motions like vibrations, short range rotations or secondary relaxations take place. At the glass transition temperature T_g the primary relaxation (alpha relaxation) becomes active allowing the system to flow. The time range over which this relaxation takes place easily covers more than ten orders of magnitudes in time. The following rubbery plateau in the modulus relates to large scale motions within a polymer chain. Two ascpects stand out. The first is the entropy driven relaxation of out of equilibrium fluctuations, secondly these relaxations are limited by confinement effects caused by the mutually interpenetrating chains. As we shall see later, this confinement is modelled most successfully in terms of the reptation model that was developed by de Gennes. Finally, when the chain has lost the memory of its confined state, liquid flow sets in. That is characterised by the translational centre of mass diffusion of the chain. Depending on the molecular weight, the characteristic length scales from the motion of a single bond to the overall chain diffusion may cover about three orders of magnitude, while the associated time scales easily may strech over more than ten orders.

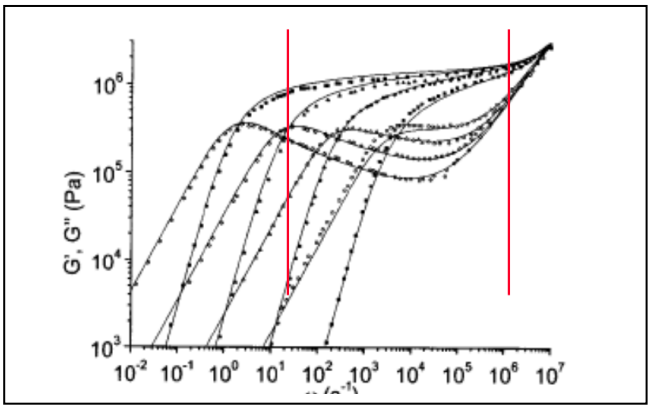


Fig. 2: *Real and imaginary part G' and G'' of the dynamic modulus for polymer melt. The two red lines display the region where the polymer melt responses elastically.*

Figure 2 quantifies this behaviour on the example of the real and imaginary part of the dynamic modulus, that is plotted as a function of frequency over about ten orders of magnitude. The parameter for the different curves is the molecular weight. The larger the molecular weight, the broader the spectrum of the modulus. Looking on the real part G' we realise a plateau in frequency that enlarges with increasing molecular weight. In this regime the polymer liquid responses elastically like a rubber. Only at low frequencies we see the transition to liquid like flow. Likewise the imaginary part G'' that describes the dissipative behaviour of the melt exhibits a maximum where the liquid flows sets in. At this point a transition from elastic to liquid like behaviour occurs. The dynamic modulus displays the viscoelastic properties of polymer melts. In a certain frequency range the elastic behaviour prevails while in others we deal with typical liquid like behaviour.

Another characteristic behaviour of polymer melts are universal power laws in the molecular weight dependence of viscosity and diffusion.

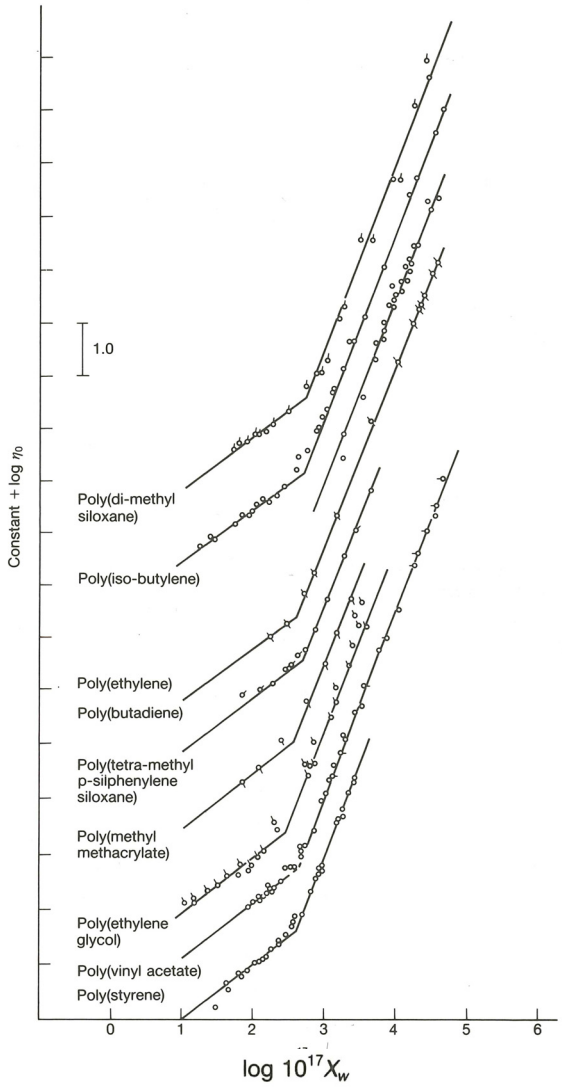


Fig. 3: *Polymer melt viscosity for various polymers as a function of molecular mass. The data are characterized by cross over between two power laws.*

Figure 3 presents the molecular weight dependence of the melt viscosity for a number of different polymers in a double logarithmic plot. In all cases the viscosity shows two different power law regimes. At low molecular weight the viscosity increases proportional to molecular weight, while above a critical molecular weight M_c, the viscosity increases dramatically with M following a power law with an exponent of about 3.4. Thus, with increasing chain length a polymer melt becomes very tough and viscous.

Likewise, the molecular weight dependence of the translational diffusion coefficient of a polymer in a melt is characterized by two different power law regimes.

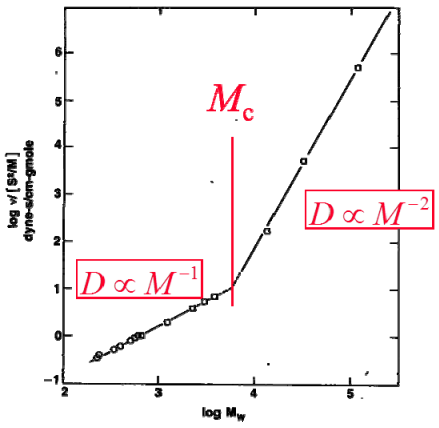


Fig. 4: *Translational diffusion coefficient in a polymer melt measured for polyethylene as a function of molecular weight: Again two different power laws are visible.*

This is shown in figure 4, where the inverse diffusion coefficient is displayed as a function of molecular weight in a double logarithmic form. Again, at low molecular weight we realise a linear dependence while at higher molecular weight the diffusion coefficient is inversely proportional to the square of the molecular weight. These earlier results by Persson at al [5]were later modified by more precise experiments reveiling a power law exponent of -2.3 instead of -2 for the molecular weight dependence of the diffusion coefficient [6].

In the following we will now ask for a molecular understanding of this peculiar behaviour of polymer melts. We will go through a hierarchy of models that will let us understand, why long chain molecules exibit the shown dynamical features.

3 The Standard Model of Polymerdynamics (Rouse)

If we want to describe the motion of a polymer, we could start with the atoms of a chain and solve Newton's equations. This asks us to deal with very many variables - already the simplest polymer chain polyethylene built from CH_2 units, at a reasonable length of about a thousand units features already 3 000 atoms. A melt of such chains gets difficult to treat already for advanced molecular dynamics simulations. We may make a step further and coarse grain in a way, that we describe the atoms along one bond, in this case the CH_2 unit by one entity leading to the unified atom model. In this case, we still have thousand atoms in one chain. Again we need severe MD simulation in order to solve the problem and we still don't have a model. In order to go further, we have to coarse grain significantly more and still keep the essentials of the problem (Fig. 5).

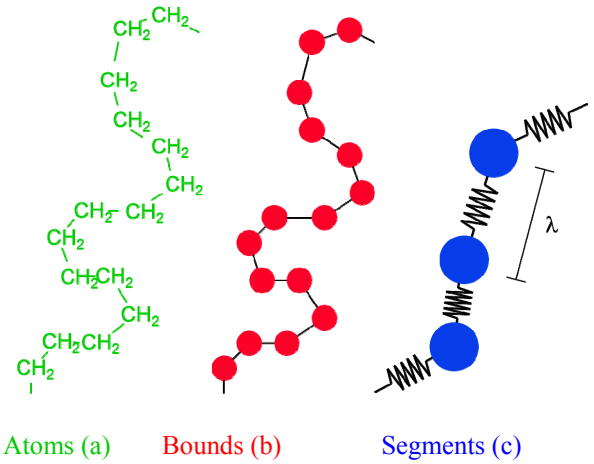


Fig. 5: *Schematic representation of the coarse graining process that leads to the Rouse model.*

a) Polyethylene chain
b) Unified atom chain
c) Chain represented by beads and springs

This is achieved with the Rouse model [7]. Here the polymer chain is described by a sequence of beads and springs where the beads undergo friction with a heat bath. The springs originate from the chain entropy that prefers a Gaussian chain conformation. Any deviations from such conformations undergo a restoring force of harmonic character.

With that we can write down a Langevin equation for the chain segmental motion.

$$\zeta_0 \frac{dr_n}{dt} = \frac{k_B T}{l^2}(r_{n+1} - 2r_n + r_{n-1}) + f_n(t) \tag{1}$$

Here ζ_0 is the friction of the beat with the heat bath $k_B T/l^2$ is the entropic spring constant, where l is the segment length and $f_n(t)$ describes the thermal random force acting on bead "n". Assuming white noise this equation can be solved exactly (see appendix). For the viscosity of the melt we get

$$\eta = \frac{\zeta_0 l^2 N_\rho}{36} \tag{2}$$

where ρ is the density and N the chain length. Note, that this viscosity is proportional to the number of chain segments N.

Similarly for the translational centre of mass diffusion coefficient Eq. [1] leads to.

$$D = \frac{k_B T}{\zeta_o N} \propto N^{-1} \tag{3}$$

The diffusion coefficient is inversely proportional to the number of friction exerting beads. For short chains both results agree with macroscopic experiments, but we may ask is this model correct microscopically?
For this purpose we have to look on the chain motion on the scale of the chain. There a prominent quantity is the mean square segment displacement. For long times we know that the motion has to be diffusive and therefore $< \Delta r_n^2(t) >= 6Dt$.

But what happens, when the chain segments have moved distances smaller than the chain dimensions? Let's start with a borderline case, the case where the covered distance is just the chain size. This obviously is the longest time, where internal correlations within the chain could play a role. This longest relaxation time is also called the Rouse time τ_R. From the simple consideration we would get $6D\tau_R = R^2 = Nl^2$.

Using Eq [3] and equating for τ_R we get

$$\tau_R \simeq \frac{N^2 l^2 \zeta_0}{6k_B T} \tag{4}$$

differing only slightly from the correct value. A solution of Eq.[1] yields $\tau_R = \frac{N^2 l^2 \zeta_0}{3\pi^2 k_B T} = \tau_0 N^2; \tau_0 = \frac{l^2 \zeta_0}{3\pi^2 k_B T} = \frac{1}{W\pi^2}$. Now, any subsection of the chain with $\frac{N}{p}$ segments relaxes as the whole chain. Therefore, we have

$$\tau_p \propto \tau_0 \left(\frac{N}{p}\right)^2 \tag{5}$$

at the time τ_p the chain section with N/p monomers moves over its own distance. Thus,

$$< (r_n(\tau_p) - r_n(0))^2 >= l^2 \frac{N}{p} = l^2 \left(\frac{\tau_p}{\tau_0}\right)^{\frac{1}{2}} \tag{6}$$

the most right part of Eq. [6] follows from Eq. [5]. Since this is true for all p, the chain segments at times shorter than τ_R move in a subdiffusive way. The mean square displacement only increases with the square root of time. This is a basic prediction of the Rouse model. Performing the full calculation starting from Eq. [1], the final result for the time dependent mean square displacement is:

$$< \Delta r^2{}_n(t) >= \sqrt{\frac{12 l^2 k_B T}{\pi \zeta_0} t} \tag{7}$$

The segment self correlation function that is measured with quasielastic incoherent neutron scattering directly accesses this quantity. In Gaussian approximation we have

$$S_{self}(Q,t) = \exp\left[-\frac{Q^2}{6} < r^2(t) >\right] = \exp[-D_R Q^2 t] * \exp\left[-\frac{2}{\sqrt{\pi}} (\Omega_R(Q)t)^{\frac{1}{2}}\right]; \Omega_R(Q) = \frac{k_B T l^2}{12\zeta_0} Q^4 \tag{8}$$

The second part of this equation is obtained by inserting Eq. [7]. $\Omega_R(Q)$ is the characteristic relaxation rate, that increases with the momentum transfer Q^4.

Even though a clear cut prediction, experimentally the observation of the self correlation function of a Rouse chain is an important challenge. The necessary resolution at the low

momentum transfers requires, neutron spin echo spectroscopy [8]. Here, incoherent experiments are difficult, since incoherent scattering depolarises the beam to a large extend (2/3 spin flip scattering). Therefore, the first successful experiments were carried out by using a trick. The chemists produced deuterated PDMS where randomly short protonated sections were copolymerised. These protonated sections in a generally deuterated environment gave rise to coherent scattering, however, since the scattering from different labels was uncorrelated the self correlation function was measured.

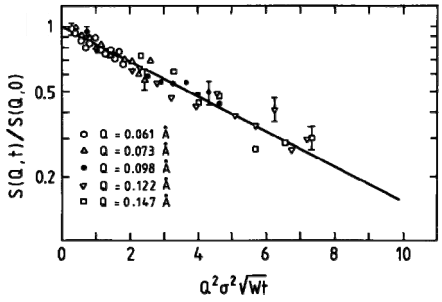


Fig. 6: *Dynamic structure factor for the segmental self motion in a PDMS-melt. The data are scaled with the Rouse variable. Solid line predicted $\sqrt{t}$ relaxation by the Rouse model.*

Figure 6 displays the obtained self correlation function for PDMS in a presentation where the logarithm of the scattering function is plotted versus $(\Omega_R(Q)*t)^{1/2}$ [^9] the scaling variable of Eq. [8]. In this way all the data collapse on one single master curve, that according to Eq. [8] should be a straight line. The experimental results beautifully verify the major prediction of the Rouse model and show that the simple approximation of the bead - spring model properly accounts for the segmental dynamics of the PDMS chain on the space time frame investigated.

For the single chain dynamic structure factor, where we look on a labelled e. g. protonated chain in a deuterated environment, we have to deal with the interference of scattered waves originating from the different atoms or monomers of the chain. The detailed calculations are displayed in the appendix. The main result is, however, similar to that for the self correlation function. The chain still follows the universal decay if plotted versus the Rouse variable $(\Omega_R*t)^{1/2}$.

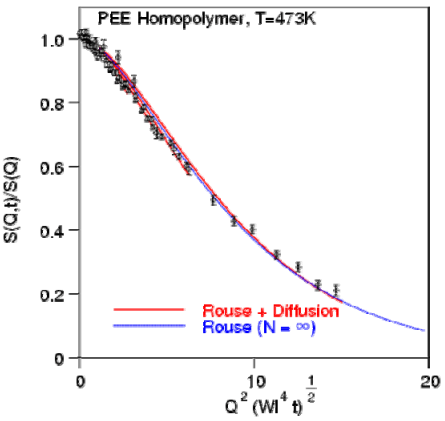


Fig. 7: *Single chain dynamic structure factor measured for polyethylethylene scaled with the Rouse variable. All data collapse to a single master curve (see text).*

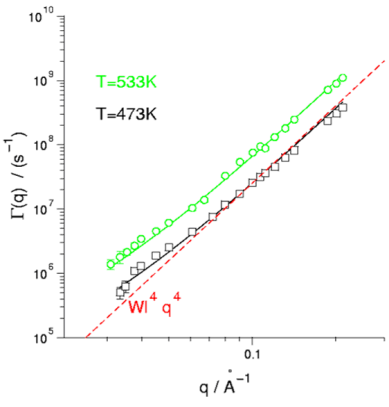


Fig. 8: *Characteristic relaxation rate for polyethylethylene at two different temperatures as a function of momentum transfer Q. The data display the predicted Q^4 scaling. For the deviations at low Q see text.*

This is displayed for polyethylethylene (PEE) in figure 7 [10] again the data follow closely a single master curve, however, with a shape that differs from the simple form of figure 6. Furthermore, again the characteristic rate Ω_R is predicted to follow the Q^4 power law.

This is beautifully seen for polyethylethylene in figure 8, where the characteristic width $\Gamma(g)$ is displayed as a function of Q in a double logarithmic way. For more than three orders of magnitude in rate, the data display the Q^4 law. The small deviations at low Q result from the the translational diffusion coefficient, that additionally contributes to the overall relaxation. From Eq. [8] follows $\Gamma(Q) = D_R Q^2 + \Omega_R(Q) = Q^2(D_R + \frac{k_B T l^2 Q^2}{12\zeta_0})$.
The solid lines in figure 8 reflect the combined diffusive and internal relaxation behaviour.

In summary the chain dynamics at intermediate scales is excellently described by the bead-spring Rouse model

- it predicts correctly, the Q-dependence of the relaxation rate (Rouse scaling variable)
- it provides quantitative expressions for the spectral shape of self and pair correlation functions
- it furthermore establishes the correct reference to macroscopic quantities as viscosity ζ and translational diffusion coefficients D_R

4 Long Chains: Entanglements

Macroscopically the dynamics of long chain polymer melts is characterised by a plateau regime in the dynamic modulus. Thus, there is a frequency or time regime where a polymer melt responses elastically like a rubber. There, the elastic properties are derived from the entropy elasticity of the chains between permanent cross links. The modulus of a rubber is inversely proportional to the mesh size and proportional to the temperature. In analogy it is suggestive to assume that in a polymer melt entanglements or topological interactions between chains take the role of the rubber cross links. They are supposed to form a temporary network, which displays the rubber elastic properties. However, other than in a rubber, for long times the chains may disentangle and the melt flows. Therefore, the dynamic modulus decays for long times or low frequencies. Using the analogy to the modulus of a rubber, we may estimate the distance between entanglement points from the value of the modulus associated to the plateau, the plateau modulus G_N^0. For different polymers these distances come out to be between about 30 and 100 Å. On that basis a number of theories for viscoelasticity have been developed. The most famous of them is the reptation model by de Gennes [3] and Doi and Edwards [2]. In this model the dominating chain motion is the reptile like creep along the chain profile. The lateral restrictions by the interpenetrating chains are modelled by a tube of size “d” parallel to the coarse grained chain profile. According to theory d relates to the plateau modulus of the melt.

$$d^2 = \frac{4}{5}\frac{R_E^2}{M}\frac{k_B T}{G_N^O} \tag{9}$$

Where R_E is the chain end to end distance and M its mass. The width of the tube or the distance in between entanglements is a mesoscopic quantity significantly larger than the distance between the chain back bones. Thus, there is lateral freedom on intermediate scales and only large scale motion is affected by the tube constraints.

These lateral constraints were first visualized in a famous computer simulation approach by Kremer and Grest [11] in 1990.

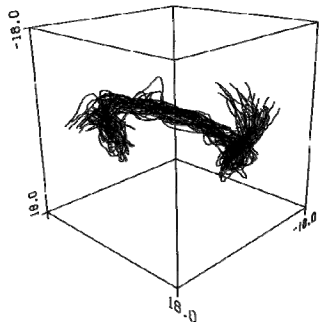


Fig. 9: *Results for the chain contour from a computer simulation by Kremer and Grest [11]. The different lines represent the chain contours at different times.*

Figure 9 displays their results. What is seen are contour plots of a polymer chain in a melt at different times. We clearly see, that in the middle part of the polymer the chain trajectory stays constrained in a tube like fashion. On both ends however, the chains are able to leave the constraints and explore new areas in space. This is the essence of the tube model.

The tube model also makes the main macroscopic finding for the molecular weight dependence of viscosity and translational diffusion comprehensible. The viscosity relates to the longest relaxation time in a system. If we consider Rouse diffusion along the tube with a Rouse diffusion coefficient $D_R \approx 1/(N\zeta_0)$ then an initial tube configuration is completely forgotten when the mean square displacement along the tube $\langle r^2(t)\rangle_{\text{tube}}$ = (contour length L)^2. Thus, for the longest relaxation time we obtain

$$\tau_\eta \approx \frac{L^2}{D_R} \approx N^3\zeta_0 \tag{10}$$

The diffusion coefficient is found by considering that during this time in real space the MSD just amounts to the end to end distance of the chain squared. Thus, we obtain

$$D_{rep} = \frac{R_E^2}{\tau_\eta} \approx \frac{N}{N^3\zeta_0} \approx \frac{1}{N^2\zeta_0} \tag{11}$$

4.1 Mean square displacements

We now consider the predictions of the reptation model for the mean square displacements of the chain segments [12,13]. For short times, when the chain segments have not yet realized the topological constraints ($r^2 < d^2$) we expect unrestricted Rouse motion $\langle r^2(t)\rangle \approx t^{1/2}$ (Eq.[7]).

Experimentally this was the case for PDMS (Figure 6). At a time $\tau_e = \frac{\pi}{4}\tau_o N_e^{\,2}$, where N_e is the length of an entanglement strand, the mean square displacement reaches the order of the tube diameter. τ_e is derived as the Rouse time for a polymer strand, spanning the tube. Thereafter motional restrictions are expected.

For times $t > \tau_e$ one dimensional curve linear Rouse motion along the tube needs to be considered. Displacements along the tube are described by Eq.[7] where we have to change real space coordinates to coordinates $s(t)$ along the tube. If a segment is displaced along the tube by $\langle (s_n(t) - s_n(0))^2 \rangle$ then the mean square displacement in 3-d real space is $d\left\langle \left(s_n(t) - s_n(0)\right)^2 \right\rangle^{1/2}$. With that we obtain

$$\left\langle r^2(t) \right\rangle = \begin{cases} 2d\left(\dfrac{k_B T \ell^2 t}{\zeta_0 \pi}\right)^{1/4} & \tau_e < t < \tau_R \\ \\ 2d\left(\dfrac{k_B T t}{N\zeta_0}\right)^{1/2} & \tau_R < t < \tau_d \end{cases} \qquad (12)$$

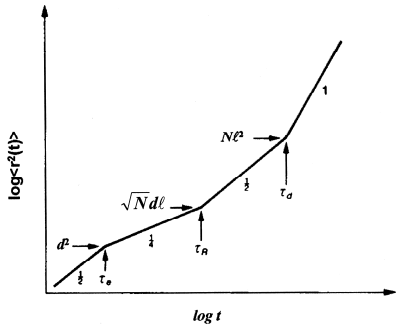


Fig. 10: *Predicted time dependence of the segmental mean square displacements. Several power laws are visible. At short times we observe the Rouse regime. (proportional $t^{1/2}$) then local reptation takes over (proportional $t^{1/4}$) there after reptation prevails ($t^{1/2}$) and finally translational chain diffusion occurs (proportional t).*

In Figure 10 the two situations correspond to the second and the third process. The second process where the chain performs Rouse motion along the tube is called local reptation while the creep like diffusion along the tube, that eventually leads to a complete tube renewal is also termed pure reptation. The terminal time τ_d after which the chain has left its original tube determines to a large extend the viscosity of the melt ($\tau_d \approx \tau_\eta$ see Eq.[10]). Beyond that time reptation diffusion prevails.

As we have alluded to in context with the Rouse Model, in Gaussian approximation the self correlation function directly relates to the mean square displacement. If the Gaussian approximation would be valid, we would assume that the width of the Gaussian function after a diffusion time t of a single segment distribution along the one dimensional tube contour may be taken as a time dependent displacement. However, as Fatkullin and Kimmich [14] have shown, the real process has to be modelled by projecting the segment probability distribution due to the Rouse motion along the random walk like contour path of the tube, this leads to a non Gaussian probability distribution of the segments at times $t > \tau_e$.

$$S_{self}(Q,t>\tau_e)=\exp\left[\frac{Q^4d^2}{72}\frac{\langle r^2(t)\rangle}{3}\right]erfc\left[\frac{Q^2d}{6\sqrt{2}}\sqrt{\frac{\langle r^2(t)\rangle}{3}}\right] \tag{13}$$

Thus, the Gaussian approximation for times longer than τ_e is invalid. The effect on the scattering function is, if it is wrongly interpreted in terms of the Gaussian approximation, that the cross over to local reptation appears to occur at a significantly shorter time than τ_e. However, the generic asymptotic behaviour remains untouched. In the sense of Eq. [13] the mean square displacement of a chain segment may still be directly observed by incoherent quasielastic neutron scattering. In the local reptation regime, we expect to observe the predicted cross over from a $t^{\frac{1}{2}}$ to a $t^{\frac{1}{4}}$ law.

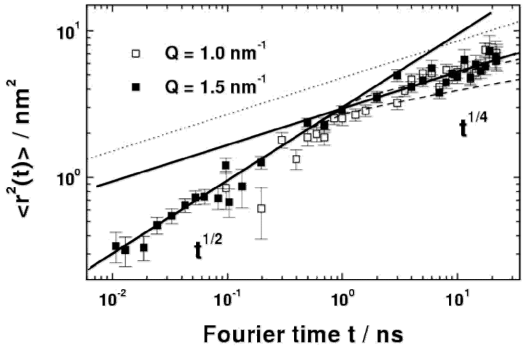


Fig. 11: *Meansquare displacements taken from a measurement of the incoherent structure factor from a fully protonated PE Melt assuming the Gaussian approximation. The solid lines describe the asymptotic power laws {$r^2(t)$} proportional $t^{1/2}$, $t^{1/4}$. Dotted lines: predictions from the Gaussian approximation; dashed lines (see text).*

Fig. 11 presents experiments performed on polyethylene samples at a temperature, where also the single chain dynamics structure factor was studied [15]. In fig. 11 the data are displayed in terms of an effective meansquare displacement, assuming the Gaussian approximation. We nicely observe the predicted cross over from $t^{\frac{1}{2}}$ to a $t^{\frac{1}{4}}$. However, if we insert the dynamic parameters for polyethylene into Eq. [12] we would expect a meansquare displacement along the tube according to the dotted line in fig. 11. The discrepancy explains itself in considering the non Gaussian character of the curve linear motion according to Eq. [13] (upper dashed line for Q=0,1Å^{-1}

4.2 Single chain dynamic structure factor

We now turn to the conceptionally more demanding question of the coherent scattering from a single labelled chain that is given, by the single chain dynamic structure factor $S_{chain}(Q,t)$.
This quantity is strongly affected by topological tube constraints.

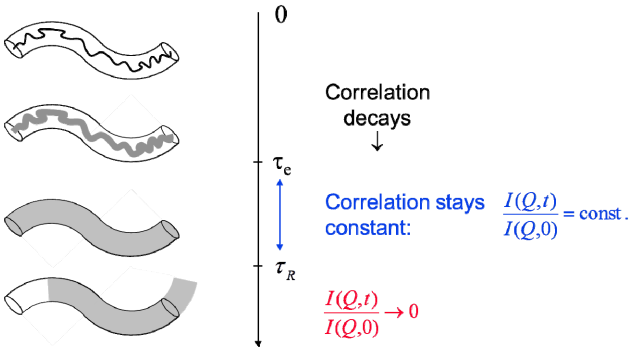


Fig. 12: *Schematic presentation of the various stages in the time development of the single chain dynamic structure factor. At short times unrestricted Rouse Dynamics time takes plac;e beyond τ_e fluctuations fill the tube; For times larger than τ_R the chain creeps out of the tube.*

Figure 12 visualises the concept, that we will now go through as a function of time.

(i) At short times $t<\tau_e$ the chain will perform unrestricted Rouse motion and the dynamic structure factor for Rouse motion (A.14) should well describe the dynamics. This is exemplified in figure 8 for dynamic data from an entangled PEE melt. At short times the tube constrains are not yet effective. In this way the chain explores the lateral constraints set by the tube. Density fluctuations of the chain are laterally equilibrated across the tube profile.

(ii) Once, this is achieved further density fluctuations of the labelled chain will only be possible via Rouse relaxation along the tube. Under such circumstances the structure factor to a first approximation mirrors the form factor of the tube. The correlations will stay and the scattering experiments will reveal the size of the topological constraints.

(iii) In the creep regime $t>\tau_R$ the memory of the tube confinement will be gradually lost and the dynamic structure factor should reveal the fraction of the still confined polymer segments.

(iv) Finally in the diffusive regime the chain reptation diffusion coefficient will be measured.

De Gennes [16] and Doi and Edwards [12] have formulated tractable analytic expressions for the dynamic structure factor. Thereby, they neglected the initial Rouse regime i.e. the derived expression is valid only for $t > \tau_e$ once confinement effects become important. The dynamic structure factor is composed from two contributions S^{loc} and S^{esc} reflecting local reptation and escape processes from the tube.

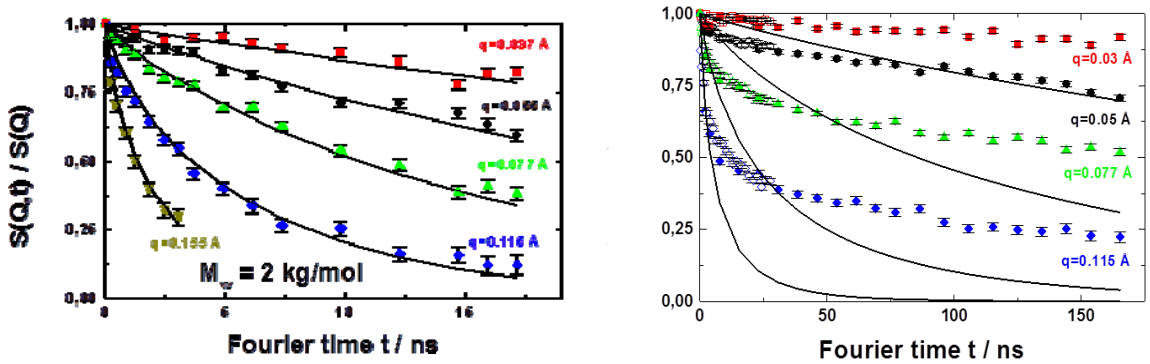
$$\frac{S_{chain}(Q,t)}{S_{chain}(Q)} = \left[1-\exp\left(-\frac{Q^2d^2}{36}\right)\right] S^{loc} + \exp\left(-\frac{Q^2d^2}{36}\right) S^{esc} \qquad (14)$$

The local reptation part was calculated as

$$S^{loc}(Q,t) = \exp\left(\frac{t}{\tau_0}\right) erfc\left(\sqrt{\frac{t}{\tau_0}}\right) \qquad (15)$$

with $\tau_0 = \dfrac{36}{Wl^4Q^4}$

A general expression for $S_{esc}(Q,t)$ due to pure reptation was given by Doi and Edwards [12]. For short times $S_{chain}(Q,t)$ decays mainly due to local reptation (first term) while for longer times (and low Q) the second term resulting from the creep motion is important. The ratio of the two relevant time scales τ_0 and τ_d is proportional to N^3. Therefore, for long chains at intermediate times a pronounced plateau in $S_{chain}(Q,t)$ is predicted. Such a plateau is a generic signature for confined motion.



(a) (b)

Fig. 13: *Single chain dynamic structure factors from PE melts at 509K (a) M_w=2000; (b) M_w=12400). The solid lines display the predictions of the Rouse Model.*

In order to illustrate the effect of topological constraints of the tube, figure 13 compares experimental results from two different polyethylene (PE) melts both studied at 509K. Figure 13a displays the results for the short chain melt (M_w=2kg/mol) The solid lines show the prediction of the Rouse dynamic structure factor (Eq. [A14])- very good agreement is achieved. Figure 13b presents equivalent results from the higher M_w melt (M_w=12.4kg/mol). The solid lines again show the predictions of the Rouse model. Also note, that the time scale for the higher molecular weight sample is extended by one order of magnitude. For the long chains the Rouse model fails completely. Only in the short time regime the initial decay of the structure factor is depicted, while for longer times the relaxation behaviour is strongly retarded signifying the tube effects.

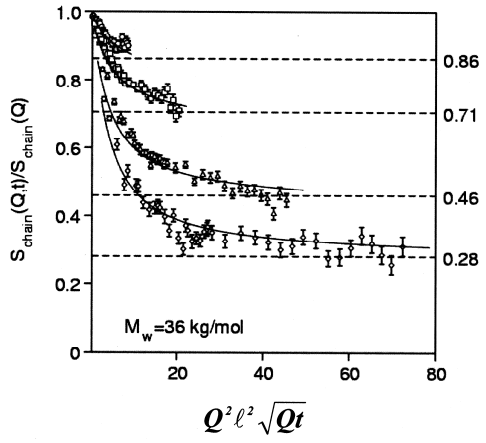


Fig. 14: *Single chain dynamic structure factor from a long polyethylene melt scaled with a Rouse variable. The dashed line described the Debye Waller factor approximation for the long time plateaus (see text).*

Figure 14 presents the dynamic structure factor from M_w=36kg/mol PE melt. Now as a function of the Rouse variable $(\Omega_R * t)^{1/2}$ [17] (see Eq. [8]) other than in figure 8, where the scaled data followed a common master curve, here, the data split into different branches that

only at a small value of the scaling variable are coming close together. The fact that the data do not follow the common Rouse scaling is the direct consequence of the dynamics length scale of the tube, that invalidates the Rouse scaling properties. We note, that this length is of dynamical character and can not be observed in static equilibrium experiments.. The heights of the achieved plateaus allow a first estimate for the amount of confinement. If we identify the plateau levels with a Debye Waller factor as a measure for the confinement, we get d=46Å, a value, that is a lower estimate for the tube diameter, since S^{loc} is not fully relaxed. The horizontal lines in figure 14 are the predictions from this Debye Waller factor estimate.

4.3 Limits of the Reptation Model

If we compare the predictions of simple reptation for the molecular weight dependences of viscosity and diffusion coefficient with the experimental findings displayed in chapter 2 we realise, that the predictions are only qualitatively in agreement with theory:

(i) Experimentally the viscosity is found to generally follow a power law proportional to $M^{3.4}$ instead of M^3.

(ii) The translational diffusion coefficient is found to behave as $M^{-2.3}$ instead of M^{-2} required by reptation.

(iii) Also the detailed frequency dependence of the dynamic loss modulus G'' does not follow the predicted $\omega^{-\frac{1}{2}}$ behaviour but rather is found to display a $\omega^{-\frac{1}{4}}$ law .

In order to cure the short comings a number of additional relaxation processes were introduced, that are consistent with reptation. The most prominent among them are contour length fluctuations (CLF) and constraint release (CR). The CLF effect evolves from the participation of the chain ends in the local reptation process (figure 15).

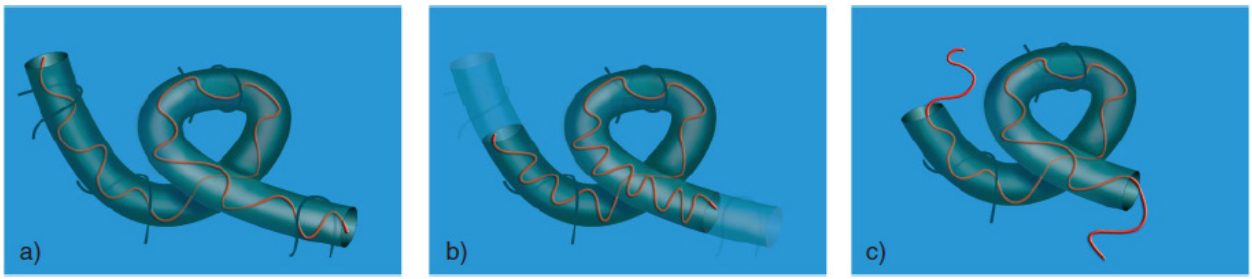


Fig. 15: *Principles of contour length fluctuations*
a) The surrounding chains confine a given chain inside a virtual tube of diameter d
b) The chain ends perform fluctuations retracting into the original tube; thus, forgetting the initial confinement of the vacated tube parts
c) The chain ends may then explore new regions in the melt

Any chain retraction and subsequent expansion will lead to a loss of memory of the original tube confinement. Thus, the tube becomes shorter with time. Mathematically the problem is treated as a first passage problem. Whenever a tube contour s is visited by a free end the tube

ceases to exist. The functional form of the tube survival probability μ(t) was derived theoretically as [18]

$$\mu(t) = 1 - \frac{1.5}{Z}\left(\frac{t}{\tau_0}\right)^{\frac{1}{4}} \qquad (16)$$

where Z is the number of entanglements along the chain. Eq. [16] provides quantitative knowledge on the chain fraction that at a time t is still confined. For polyethylene all parameters are known from NSE experiments on asymptotically long chains where the CLF effect does not play a role.

With this knowledge an experiment was designed [19] where the dynamic structure factor of a chain that is subject to CLF was compared with that of an identical chain where the contrast of those segments which within the experimental time frame are affected by CLF was matched.

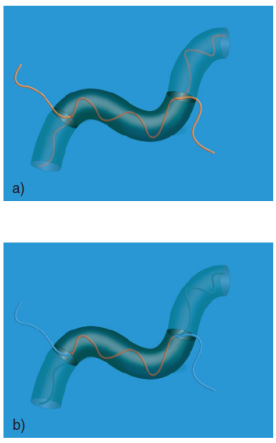


Fig. 16:

a) For the fully protonated chain in a deuterated matrix local reptation and contour lengths fluctuations are visible

b) For a center labelled (protonated chain with deuterated chain ends) in a deuterated matrix only local reptation is visible. The contour lengths fluctuations of the chain ends are masked out.

The experimental idea is displayed in Fig. 16: case (a) is realised by performing an experiment on a fully protonated chain in a deuterated matrix. In this case, the full chain dynamics including the CLF is observed. Case (b) is realised by a chain, where the inner part is protonated and the two outer chain sections of a length, that is affected by CLF are deuterated and therefore not visible in the deuterated matrix. In such a case the dynamics of the chain should be equal to that of an asymptotically long fully confined chain.

With the known parameters for PE Eq. [16] reveals that on average on each side 220 monomers are released during the observation time of 190 ns at 509 K. The experiments were performed on two different chains of a molecular weight of 25kg/mol one was fully hydrogenated and the other had deuterated labels of about M_w=4kg/mol corresponding to 260 monomers on each end. Both were studied in a deuterated matrix of the same molecular weight.

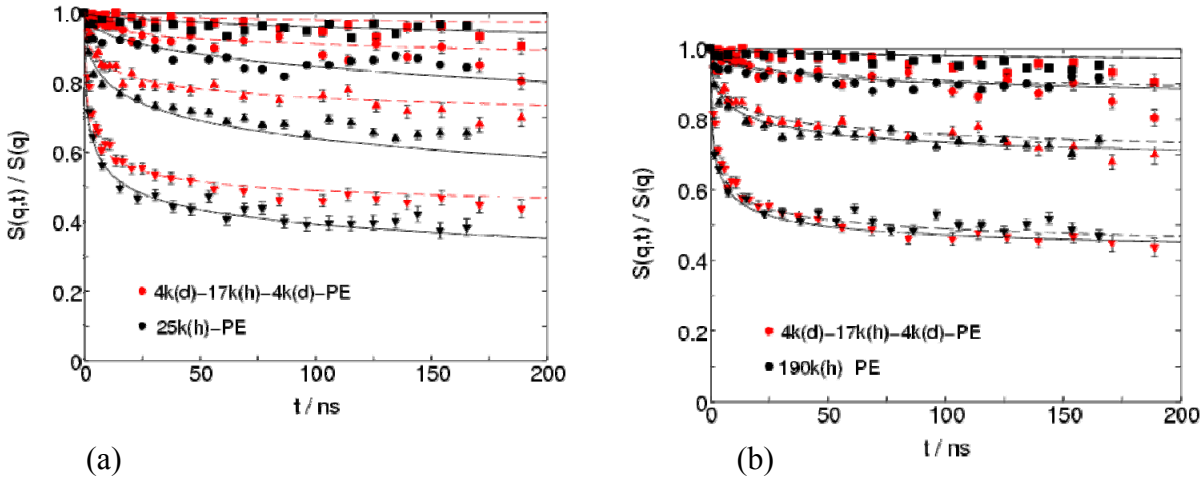


Fig. 17: *(a) Comparison of the single chain dynamic structure factors from an end labelled chain and a fully labelled chain. The solid and dashed lines represents fit with the corresponding models.*
(b) Comparison of the single chain dynamic structure factors for a very long and an end labelled polyethylene chain. The inner part of the shorter end labelled chain relaxes as the very long chain.

Figure 17 presents the normalized dynamic structure factors S(Q,t)/S(Q) as a function of time for different Q values. 17 (a) compares the results for the differently labelled chains. The fully labelled chain relaxes significantly stronger than the corresponding center labelled counter part. Apparently the constraints are stronger for the center labelled chain than for the chain where the outer parts are also visible. We also note, that in the case where the ends were masked the chain center part shows exactly the same structure factor as that for a very long chain where CLF processes do not play a role (figure 17b). This result directly demonstrates that the action of CLF happens at the chain ends and that the confinement remains fully conserved with time in the center. The agreements of the data sets in figure 17b also show, that constraint release in the investigated space time regime does not play any role. The solid lines in the two figures represents theoretical descriptions of the CLF effect and are taken from reference [17].

At the end we like to emphasize, that CLF also affects the macroscopic melt properties in a significant way. (i) It has been shown, that CLF introduce the $\omega^{-\frac{1}{4}}$ regime into the spectrum of G''(ω). However, the other limiting mechanism (CR) adds further significant modifications. (ii) The translational chain diffusion necessarily is also affected. First, the terminal time defining the diffusion step is reduced since reptation has only to relax the not yet released central parts of the tube and secondly the diffusive lengths is reduced, since only the displacement of the central part counts. Both effects do not cancel and the net effect gives D proportional $(N/N_e)^{-2.4}$ in very good agreement with experiments. (iii) Finally also the anomalous power law exponent of 3.4 for the viscosity mass relation has been attributed to CLF.

It is the virtue of the NSE experiments, that they confirm the CLF mechanism quantitatively on a molecular level in space and time. For similar experiments on the constraint release mechanism, we refer to reference [20].

5 Functional dynamics of proteins

While the study of the dynamics of synthetic polymers has reached some maturity, the next challenge is the investigation of the large scale motion of biopolymers. The goal there is to find out to what extend these dynamics play a role in bio function. On local scales some insight into the conformational dynamics has been gained e.g. by time dependent crystallography [21]. On the other hand, the large scale dynamics such as protein domain motions is rarely investigated experimentally, because of the lack of techniques to study these large scale correlated motions. In this lecture we present a first study on such dynamics on the example of Alcohol Dehydrogenase (ADH) [22]. The alcohol dehydrogenases are enzymes that are important for many organisms allowing the interconversion between alcohols and ketones. In humans ADH is present as a dimer and catalyzes the oxidation of ethanol allowing thereby the consumption of alcohol in beverages. In yeast on the hand it is at the basis of the fermentation process converting acetaldehyde into ethanol. In the process the cofactor Nicotinamide Adenine Dinucleotide (NAD) is needed assisting the oxidation reaction at the zinc catalytic site.

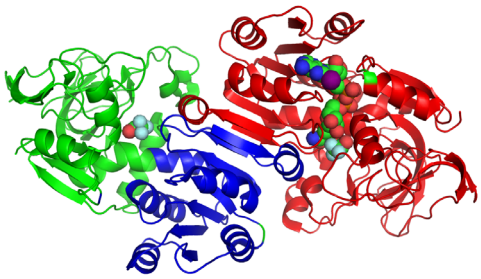


Fig. 18: *Dimer of alcohol dehydrogenase. The molecule presented by spherical caps is the NAD cofactor used in the chemical reaction.*

Based on crystallographic data Figure 18 displays a schematic structure of the dimer. The two monomeric units are clearly visible. Each monomer is build from two domains, the catalytic and the binding domain with a small opening in between. For function the small cleft between the two domains needs to open in order to allow the cofactor NAD and the ethanol to reach the catalytic Zn-atoms at the bottom of the cleft. ADH from yeast assembles in in a tetrameric structure. The crystallographic data suggest a crossed arrangement of the two dimers.

In order to verify whether in solution a similar tetrameric aggregate is present it is important to perform neutron small angle scattering (SANS) experiments. Figure 19 presents SANS data at different concentrations [20]. Let us commence with the insert.

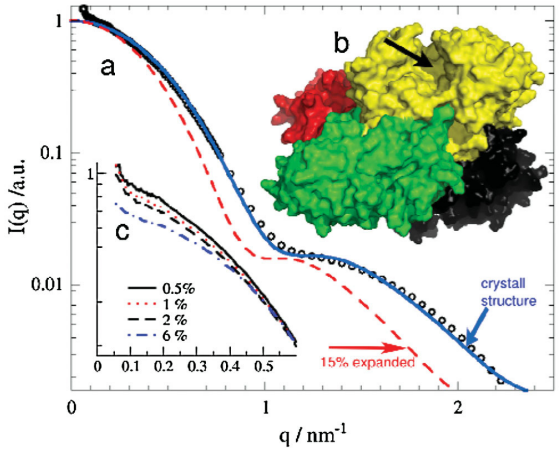


Fig. 19: *SANS results on ADH solutions of different concentration). Insert: concentration dependent results. The solid line in the main figure displays the SANS prediction on the basis of the crystal structure.*

Here for different concentrations the low Q data are presented and there with increasing concentration we observe a decrease of the SANS intensities. This observation results from the repulsive interaction between different tetramers giving rise to an inter tetramer structure factor causing the reduced intensity at low Q.

In the main figure the low concentration data are presented over the full Q range and are compared to the crystal structure that appears to be in a very good agreement with the SANS data.

Now we turn to the dynamics. Such a protein complex in solution performs various important motions: (i) translational diffusion, where the whole molecule migrates the solvent (ii) rigid body rotational diffusion. There the complex rotates around its center of mass giving rise to important dynamic contribution to the scattering signal as soon as the momentum transfer reaches the typical inverse size of the complex and finally (iii) there might be contributions from a possible cleft opening dynamics that would also be expected in a Q-range where the rotational motion contributes. Thus, in order to identify any large scale internal protein dynamics, one has to sort out all the other important motions.

To separate out the translational diffusion is done in an experiment at very small momentum transfer , where the molecule appears as a point like object. Dynamic light scattering is the proper tool since it investigates the overall dynamics on that length scale. Figure 20 displays light scattering results for different ADH concentrations as a function of Q^2. We note that apparently for all concentrations we observe identical translational diffusion coefficients. At 5°C it amounts to $D_{DLS} = 2.35 \pm 0.2 \times 10^{-2} \mathrm{nm}^2/\mathrm{ns}$.

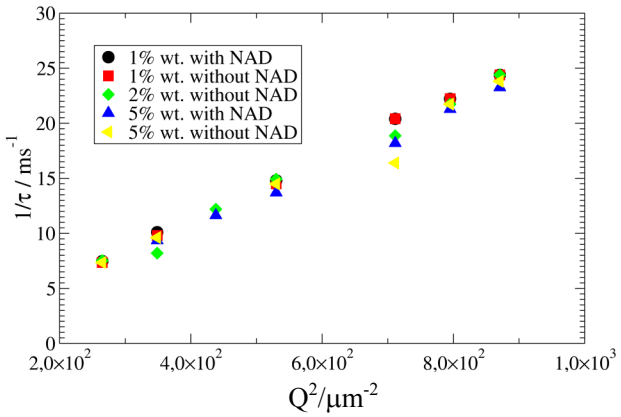


Fig. 20: *Dynamic light scattering results on the translational diffusion of tetrameric ADH in water solution at different concentrations. The characteristic rates* $1/\tau = D^*Q^2$ *are plotted vs.* Q^2.

In order to approach the internal dynamics of such an aggregate one has to increase the Q range such that $1/Q$ roughly corresponds to the molecular or aggregate size. This may be achieved implying neutron spin echo spectroscopy to the ADH solutions. Figure 21 displays NSE results for a large number of different momentum transfers Q. The data are presented in a log linear fashion showing directly the near single exponential decay observed in all cases. Fits with single exponential decays are included by straight lines. We note however, that at intermediate Q ($Q = 0.68\text{nm}^{-1}$) small but systematic deviations appear indicating a two component structure.

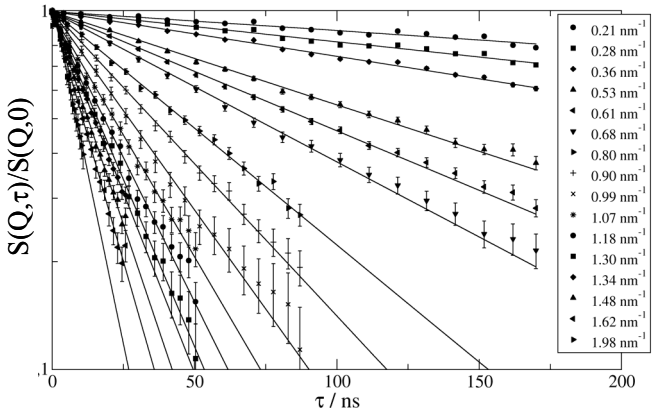


Fig. 21: *Neutron spin echo results on a 5% ADH resolution at 5°C with cofactor for various momentum transfers.*

Comparing the covered Q range with the SANS data (Figure 19) we realize that the range of the structure factor, where intermolecular interactions are important as well as the regime of internal structure are covered. In a first step we describe all data in terms of a single exponential decay. Then, we may approximate the spectra in terms of a first cumulant expansion.

$$\ell n \frac{S(Q,t)}{S(Q)} = -\Gamma(Q)t + \frac{1}{2}K_2 t^2 \qquad \text{(a)} \qquad (17)$$

The decay rate of the dynamic structure factor, also called initial slope, is

$$\Gamma(Q) = -\lim_{t\to 0}\frac{\partial}{\partial t}\,\ell n\left[S(Q,t)\right] \qquad \text{(b)} \qquad (17)$$

Using the decay rates $\Gamma(Q)$ we may define an effective diffusion coefficient

$$D_{eff}(Q) = \frac{\Gamma(Q)}{Q^2} \qquad \text{(c)} \qquad (17)$$

For the case of a translational diffusion $D_{eff}(Q)$ would be a constant and coincide with the translational diffusion coefficient.

Figure 22 displays the thus obtained effective diffusion coefficients as a function of Q for the different concentrations with and without the cofactor NAD.

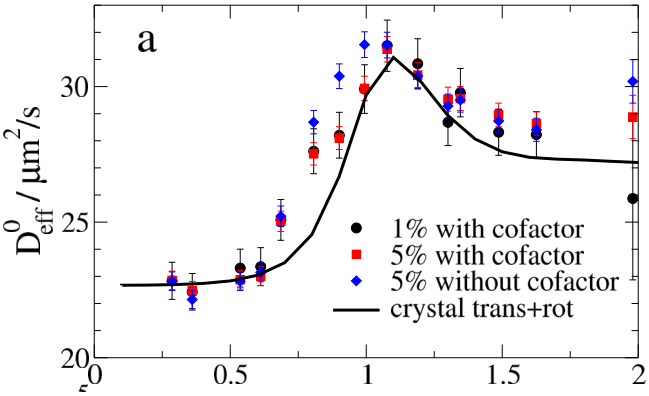


Fig. 22: *Effective diffusion coefficient $D_{eff}(Q)$ for 3 different solutions of ADH, The data were corrected for H(Q) and S(Q) (see text) Solid line results from a rigid body calculation (see text)*

The experimental results show a strong Q modulation with a maximum around $Q = 1\text{nm}^{-1}$. Furthermore, we realize that at low Q the data are in agreement with the concentration independent light scattering results. We also see that beyond the statistical error in the low Q flank of the data the relaxation without the cofactor NAD is faster than that including the cofactor. We may conclude that in the dynamics of the ADH tetramer on the scale of the aggregate itself we observe significant contributions beyond translational diffusion.

We now want to interpret the data and commence with the low Q data that are affected by the interactions between the molecules. In this regime the effective diffusion coefficient relates to the diffusion coefficient D_0 at infinite dilution by

$$D_{eff}(Q) = D_0 \frac{H(Q)}{S(Q)} \tag{18}$$

where $H(Q)$ is the hydrodynamic factor and $S(Q)$ is the interparticle structure factor. The structure factor may be extracted from the SANS data in dividing the concentration dependent results by the lowest concentration results (0.25%). With this experimental structure factor the data need to be corrected. The hydrodynamic factor cannot be measured directly. A first approximation in terms of a Perkus Yevic model shows that (i) at a 1% level the correction factor $H(Q)/S(Q)$ leaves the experimental data practically untouched and (ii) at 5% the correction is somewhat weaker than the experimentally observed effect. Nevertheless, beyond $Q = 0.6\text{nm}^{-1}$ the ratio of $H(Q)/S(Q)$ remains constant. Thus, the observed higher Q structure is entirely determined by intra aggregate effects.

The prime reason for a Q dependent structure in $D_{eff}(Q)$ are rotational motions of the molecule. In a first cumulant approximation (see Eq.[17]) the effective diffusion coefficient of a rigid boy undergoing translational and rotational diffusion has the form

$$D_{eff}(Q) = \frac{k_B T}{Q^2} \times \frac{\sum_{jk} \left\langle b_j e^{i\underline{Q}\underline{r}_j} \begin{pmatrix} \underline{Q} \\ \underline{Q}\times\underline{r}_j \end{pmatrix} \tilde{H} \begin{pmatrix} \underline{Q} \\ \underline{Q}\times\underline{r}_K \end{pmatrix} b_K e^{-i\underline{Q}\underline{r}_K} \right\rangle}{\sum_{jK} \left\langle b_j e^{i\underline{Q}\underline{r}_j} \; b_K e^{-i\underline{Q}\underline{r}_K} \right\rangle} \tag{19}$$

Here $\underline{r}_j$ and $\underline{r}_k$ are the atomic coordinates, b_j and b_k the corresponding neutron scattering lengths and $\tilde{H}$ the mobility tensor. The sum runs over all atoms of the molecule or molecular

aggregate and the pointed brackets indicate an spherical average. The denominator is the aggregate formfactor. The mobility matrix $\ddot{H}$ is a 6×6 tensor involving translational (T) and rotational (R) parts including a translational rotational coupling (TR). For the simplest case of an isotropic particle $T = D_{trans}/k_BT$ and $R = D_{rot}/k_BT$. The evaluation of Eq.[19] is importantly complicated by the hydrodynamic interaction between the different parts of the molecule. In the biophysical literature one finds the computer code HYDROPRO that was developed by the group around Garcia de la Torre [21]. In this code a complicated molecule is approximated by a rigid aggregate of little spheres. Its diffusional motion including the hydrodynamic interaction is then calculated by a proper superposition of the motion of the rigidly connected spheres. The calculations with HYDROPRO need as an input the crystallographic coordinates of all atoms. The solid line in Figure 22 displays the result of the HYDROPRO calculations for a rigid molecule. While the line describes the general form of the effective diffusion coefficient data reasonably well, we observe significant deviations at smaller momentum transfers. These differences between the rotational diffusion expectation for a rigid aggregate and the experimental data are displayed in Figure 23. We note that at Q values below the peak of the rigid body rotational diffusion coefficient significantly faster effective diffusion takes place. This result indicates the presence of internal motion within the molecule which must involve mainly those atoms which are placed in the outer regions of the tetramer emphasizing more strongly the larger distances and within the molecule and therefore giving rise to extra dynamics at low Q.

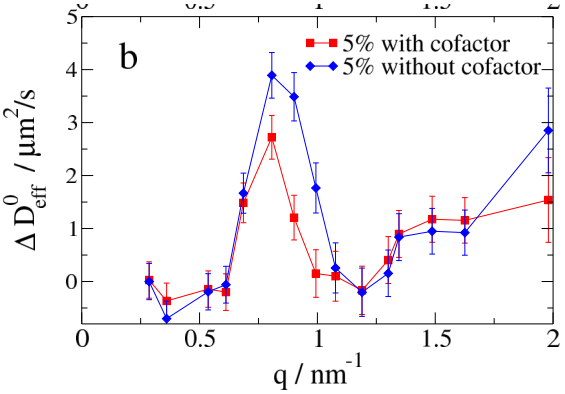


Fig. 23: *Differences between the measured effective Q- dependent diffusion coefficient and expectation from rigid body motion. The lines are guides to the eye.*

A first interpretation of this result may be carried out in terms of a normal mode analysis. For this purpose an elastic network model is used where the complicated bonded and non-bonded interactions are replaced by a pair wise Hookian potential controlled by a single parameter. As has been shown, such a simple formulation is sufficient to describe the low frequency motion of large proteins [22].

In this model two atoms are connected by a spring, if the equilibrium distance between them is sufficiently small. The elastic forces acting on the particles obey Hookes law and depend only on the change in the distances between them. If $\underline{R}^0_i$ are the equilibrium positions of the particles and $\underline{R}_i(t)$ are their actual coordinates the Hookean Potential has the form

$$E_H = f\sum_{i,j} A_{ij}\left(\left|\underline{R}_i - \underline{R}_j\right| - \left|\underline{R}^0_i - \underline{R}^0_j\right|\right)^2 \qquad (20)$$

A is the adjacency matrix with the elements, $A_{ij} = 1$ if $\left|\underline{R}_i^0 - \underline{R}_j^0\right| < \ell_0$ and $A_{ij} = 0$ otherwise. In principle the dynamics of such an elastic network is nonlinear because the distances $\left|\underline{R}_i - \underline{R}_j\right|$ are nonlinear functions of the coordinates. Close to the equilibrium the equations of motions can however be linearized yielding an equation of motion

$$m_i \ddot{\underline{r}}_i = -\sum_j \Lambda_{ij}\, \underline{r}_j \tag{21}$$

where Λ is the $3N \times 3N$ dynamical matrix obtained by the linearization process. In this linear approximation the motions are described by a sum of independent oscillating normal modes. They are represented by the non zero eigenvalues ω_α and respective eigenvectors e^α of the matrix $\mathbf{\Lambda}$.

The large scale slow motions we are interested in, are dominated by the soft modes with small eigenvalues. In the case of overdamped modes which are seen in the experiment the oscillating part needs to be replaced by an exponential $\left(e^{(-\lambda_\alpha t)}\right)$. The relaxation rates λ_α contain the unknown friction factors within the molecule and with the surrounding water molecules.

A first approximation for the dynamic structure factor may be obtained in terms of a one phonon approximation of the cross section. In this approximation the dynamic structure factor reads

$$\begin{aligned} S(Q,t)_{1-phon} &\simeq \sum_{k,\ell} b_k\, b_e\; e^{i\underline{Q}\underline{r}_k} e^{-i\underline{Q}\underline{r}_e} \\ &+ \sum_\alpha \frac{k_B T}{\omega_\alpha^2} e^{-\lambda_\alpha t} \sum_{k,\ell} \frac{b_k\, b_e}{\sqrt{m_k m_l}}\; e^{i\underline{Q}\underline{r}_k} e^{-i\underline{Q}\underline{r}_e} \left(\underline{Q}\,\underline{e}_k^\alpha\right)\left(\underline{Q}\,\underline{e}_\ell^\alpha\right) \end{aligned} \tag{22}$$

In first cumulant approximation (Eq.[17]) we have to take the logarithmic derivative with respect to time at time equal zero. Subtracting the translational and rotational part we finally obtain

$$\Delta D_{eff}(Q) = \frac{\sum_\alpha \lambda_\alpha \frac{k_B T}{\omega_\alpha^2} \sum_{k,\ell} \frac{b_k\, b_l}{\sqrt{m_k m_l}}\; e^{i\underline{Q}\underline{r}_k} e^{-i\underline{Q}\underline{r}_e} \left(\underline{Q}\underline{e}_k^\alpha\right) \times \left(\underline{Q}\underline{e}_\ell^\alpha\right)}{Q^2 \sum_{k,\ell} b_k\, b_e\; e^{i\underline{Q}\underline{r}_k} e^{i\underline{Q}\underline{r}_e}} \tag{23}$$

describing the dynamic formfactor of the eigenmodes. Finally, Figure 24 presents the outcome of such a harmonic analysis for the tetrameric aggregate of ADH with and without the cofactor. In an exemplary way we present the contributions from the modes 7 and 11 with and without the cofactor. In all cases the low eigenmodes exhibit a formfactor with the strong peak around $Q = 1\text{nm}^{-1}$. Comparing with Figure 23 qualitatively the experimental observation and the results of the normal mode analysis resemble each other.

Quantitatively the observed experimental feature is shifted towards smaller Q indicating a more pronounced motion of the outer atoms. This difference is not yet fully understood but may result from the anharmonicity of the dynamics or the effect inhomogenously distributed friction within the molecule or both.

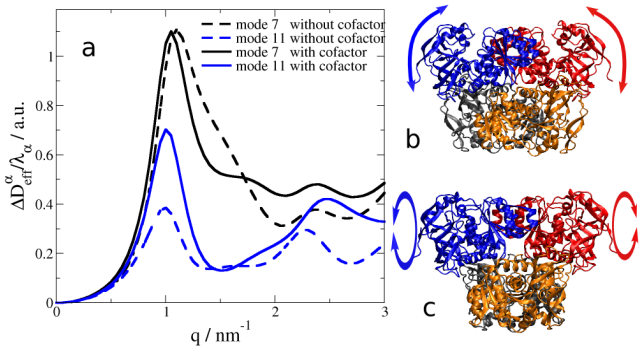


Fig. 24: a *Diffusion form factor of the normal modes 7 and 11 for the protein with and without cofactor.* ***b****, Motional pattern of mode 7: Without cofactor the exterior domain (catalytic domain) tilts outwards and opens the cleft. The inner domain with connection points between the monomers remains stiff.* ***c****, Motional pattern of mode 11: With and without bound cofactor the monomers within a dimer exhibit torsional motion around the long dimer axis (in the image plane), which is more pronounced without the cofactor.*

6 Conclusion and outlook

We have presented some representative results from neutron spin echo spectroscopy on the dynamics of macromolecules. In the case of synthetic polymers we have displayed recent results on the universal dynamics of flexible polymers from the entropy driven Rouse dynamics to confinement and reptation. In the case of biopolymers we have displayed some first experiments on the intra aggregate motion of Alcohol Dehydrogenase, an important enzyme for fermentation and the oxidation of ethanol. The lecture attempted to transmit a flavour of what can be achieved with high resolution neutron spin spectroscopy that permits access to the molecular motion simultaneously in space and time.

After some brief description of generic results on the macroscopic dynamics of polymer melts the lecture commenced with a description of the standard model of polymer motion, the entropy driven dynamics covered by the so called Rouse model. In the spatial range where the Rouse approximations are valid, the NSE measurements have confirmed most of the predictions of the Rouse model both for the self- and pair correlation function.

Towards larger scales topological interactions resulting from the mutually interpenetrating chains gain dominating influence and confine the chain motion to a tube along the chain profile. We have presented measurements on the dynamic structure factor of a reptating chain which unequivocally confirm the picture of local reptation i.e. Rouse relaxation along the contorted tube. A measurement of the self correlation function corrobates the picture. To conclude this section we have described results on contour length fluctuations that modify the reptation mechanism significantly.
Compared to the investigations of the dynamics of synthetic polymers the study of the large scale relaxation dynamics of biopolymers is still in its early stages. We have presented some first experimental data on the collective domain fluctuations of a tetrameric aggregate formed by alcohol dehydrogenase. It became possible to directly measure the Q dependent effective

diffusion coefficient which bears information on the detailed rotational diffusion dynamics. Furthermore, additional dynamics appears at low momentum transfers which is related to motions of the outer more flexible parts of each dimer. Employing normal mode analysis these dynamical features have been attributed to the slow large scale collective fluctuations of the domains.

These experiments are a first attempt to directly observe the collective internal dynamics of proteins or protein complexes. In the future experiments will be needed in order to resolve the internal dynamics of further proteins. Such experiments need to be accompanied by computer simulations in order to enhance the level of interpretation. Furthermore, the experiments need to address proteins where domain motion is functionally important. One first example are recent studies on phosphoglycerate kinase, where it could be shown that the large scale fluctuations position the substrates bound at the domains into catalytic positions. [2]

A Appendix: The Rouse Model

The dynamics of a generic linear Gaussian chain as described in the Rouse model [6] is the starting point for a description of the Brownian dynamics in polymer melts. In fact the Rouse model may be considered as the standard model for polymer dynamics and in many cases serves as a starting point for the treatment of more complex problems. In the Rouse model the conformational entropy of a chain acts as a resource for restoring forces for a chain conformation deviating from thermal equilibrium. Because of its fundamental relevance for chain dynamics problems, in this appendix we present its mathematical treatment in some detail, such that it may be followed in all steps. [1]

A.1 Gaussian chains

The conformation of a flexible linear polymer chain on scales somewhat larger than the main chain bond length ℓ_0 assumes a random walk. The conformations of such a chain are described by a set of segment vectors $\{\underline{r}(n)\} = (\underline{R}(n) - \underline{R}(n\text{-}1))$ where $\underline{R}(n)$ is the position vector of segment n. Following the central limit theorem the length distribution of a vector $\underline{r}$ connecting segments that have a topological distance of n steps is a Gaussian.

$$\phi(R,n) = \left(\frac{3}{2\pi n\ell^2}\right)^{3/2} \exp\left(-\frac{3R^2}{2n\ell^2}\right) \qquad \text{(A1)}$$

with ℓ the segment length. The Rouse model bases on a further idealization of the chain statistics assuming that the bond vector $\underline{\ell}$ of hypothetic connecting points along the chain also has a Gaussian distribution. With $\langle \ell^2 \rangle = b^2$. For simplicity throughout the lecture we will take $b^2 = \ell_0^2 C_\infty$. Keeping, however, in mind that the building block of a Gaussian chain may well contain a larger number of main chain bonds. C_∞ is the characteristic ratio accounting for the

local stiffness arising from the non-random bond angle distribution of the bonds of length ℓ_0. The probability of a conformation $\{\underline{r}(n)\}$ follows as

$$P\left(\{\underline{r}(n)\}\right) = \left(\frac{3}{2\pi\ell^2}\right)^{\frac{3}{2}N} \exp\left[-\sum_{n=1}^{N}\frac{3\left(\underline{R}(n)-\underline{R}(n-1)\right)^2}{2\ell^2}\right] \qquad \text{(A2)}$$

with N counting the number of segments of the chain. The free energy of a Gaussian chain is entirely described by its conformational entropy

$$S = k_B \ell n\left[P\left(\{r(n)\}\right)\right] \qquad \text{(A3)}$$

The Gaussian chain model yields a spring constant for each Gaussian segment. $k = 3k_BT/\ell^2$ where k_B is the Boltzmann constant. From Eq.[A2] the chain extension between arbitrary points along the chain becomes $|n - m|\,\ell^2$.

A.2 The Rouse model

The Rouse model starts from a Gaussian chain representing a coarse grained polymer model where springs stand for the entropic forces between hypothetic beads [6] (Fig. A1).

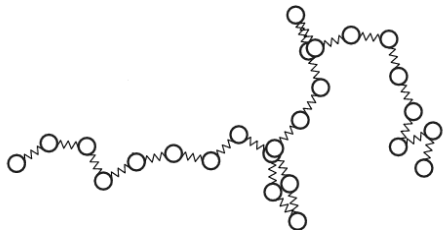


Fig. A1: *Spring-bead model of a Gaussian chain as assumed in the Rouse model. The beads are connected by "entropic springs" and are subject to a frictional force $\zeta_0\upsilon$, where υ is the bead velocity and ζ_0 the bead friction coefficient*

We are interested in the motion of segments on a length scale $\ell < r < R_E$ where $R_E^2 = N\ell^2$ is the end to end distance of the chain. The segments are subject to an entropic force resulting from the derivative of Eq.[A3] (x components).

$$\frac{\partial}{\partial x(n)}\ S = \frac{k_BT}{\ell^2}\left[x(n+1) - 2x(n) + x(n-1)\right] \qquad \text{(A4)}$$

And to a stochastic force $f_x(n,t)$ which fulfils $\langle f_x(n,t)\rangle = 0$ and $\left\langle f_\alpha(n,t)\ \ f_\beta(m,0)\right\rangle = 2k_BT\,\zeta_0\ \delta_{nm}\ \delta_{\alpha\beta}\ \delta(t)$. ζ_0 denotes the friction coefficient and α,β the Cartesian components. Regarding the index n as a continuous variable the Langevin equation for segment motion assumes the form

$$\zeta_0 \frac{\partial x}{\partial t} = \frac{3k_BT}{\ell^2} \frac{\partial x^2(n)}{\partial n^2} + f_x(n,t) \tag{A5}$$

The boundary condition of force free ends requires $\frac{\partial x(n)}{\partial n}\Big|_{n=0,N} = 0$. The partial differential equations are solved by Cosine Fourier transformation to normal coordinates fulfilling the boundary conditions. In normal coordinates the Langevin Eq.[A5] becomes

$$\zeta_p \frac{\partial}{\partial t} \tilde{x}(p,t) = -k_p \tilde{x}(p,t) + \tilde{f}(p,t) \tag{A6}$$

with $k_p = \frac{2\pi^2 p^2}{N} k$, $\zeta_p = 2N\zeta_0$. For the stochastic force we have

$$\left\langle f_\alpha(p,t)\, f_\beta(q,0) \right\rangle = 2\zeta_p\, k_BT\, \delta_{pq}\, \delta_{\alpha\beta}\, \delta(t) \tag{A7}$$

Eq.[A6] is readily solved by a single exponential

$$\tilde{x}(p,t) = \frac{1}{\zeta_p} \int_{-\infty}^{t} dt' \exp\left((t-t')/\tau_p\right) \tilde{f}(p,t') \tag{A8}$$

Where the mode relaxation time τ_p is given by

$$\tau_p = \frac{\zeta_p}{k_p} = \frac{\zeta_0 N^2 \ell^2}{3\pi^2 k_BT\, p^2} = \frac{\tau_R}{p^2} = \frac{N^2}{W\pi^2 p^2};\; W = \frac{3k_BT}{\ell^2\zeta_0} \tag{A9}$$

τ_R is the Rouse time – the longest time in the relaxation spectrum – and W is the elementary Rouse rate. The correlation function of the normal coordinates is finally obtained to

$$\left\langle \tilde{x}(p,t)\, \tilde{x}(p,0) \right\rangle = \frac{N\ell^2}{6\pi^2 p^2} \exp\left(-\frac{t}{\tau_p}\right) \tag{A10}$$

For the center of mass coordinate one finds

$$\left\langle \tilde{x}(0,t)\, \tilde{x}(0,0) \right\rangle = \frac{2k_BT}{N\zeta_0} t \tag{A11}$$

Scattering experiments relate to mean square segment correlation functions which are obtained by back transformation of the normal coordinates

$$B(n,m,t) = 3\left\langle \left(x(n,t) - x(m/o)\right)^2 \right\rangle = 6D_R t + |m-n|\ell^2 + \frac{4N\ell^2}{\pi^2} \tag{A12}$$

$$\times\sum_{p=1}^{N}\frac{1}{p^2}\cos\left(\frac{p\pi m}{N}\right)\cos\left(\frac{p\pi n}{N}\right)\left[1-\exp\left(-\frac{tp^2}{\tau_R}\right)\right]$$

In order to arrive at Eq.[A12] we have used $D_R=\frac{3}{6t}\left\langle\left(x(0,t)-x(0,0)\right)^2\right\rangle$; D_R: Rouse diffusion coefficient and $\sum_{p=1}^{\infty}\frac{1}{p^2}\left[\cos\left(\frac{p\pi m}{N}\right)-\cos\left(\frac{p\pi n}{N}\right)\right]^2=\frac{\pi^2}{2N}|n-m|$. For the special case of the self correlation function ($n = m$) $B(n,n,t)$ reveals the mean square displacement of a polymer segment. For large p the $\cos^2$ in Eq.[A12] is a rapidly oscillating function which may be replaced by the mean value ½. With this approximation we convert the sum into an integral and obtain

$$B(n,n,t)\;=\;\left\langle r^2(t)\right\rangle=2\ell^2\left(\frac{3k_BTt}{\pi\zeta_0\ell^2}\right)^{1/2}+6D_Rt \qquad \text{(A13)}$$

We find the result obtained from scaling arguments in the main text (Eq. [7]). For the translational diffusion coefficient $D_R=\frac{k_BT}{N\,\zeta_0}=\frac{W\ell^4}{3N\ell^2}=\frac{W\ell^4}{3R_E^2}$ is obtained. D_R is inversely proportional to the number of friction performing segments.

The self correlation function relates directly to the mean square displacement of the diffusing segments.

For coherent scattering observing the pair correlation function interferences from waves emanating from various segments complicate the scattering function

$$S_{chain}(Q,t)=\frac{1}{N}\exp\left[-Q^2D_Rt\right]\sum_{n,m}\exp\left\{-\frac{1}{6}|n-m|\;Q^2\ell^2\right\}\;*$$
$$\exp\left\{-\frac{2}{3}\frac{R_E^2Q^2}{\pi^2}\sum_p\frac{1}{p^2}\left\{\cos\left(\frac{p\pi m}{N}\right)\cos\left(\frac{p\pi n}{N}\right)\left(1-\exp\left(-\frac{tp^2}{\tau_R}\right)\right)\right\}\right\} \qquad \text{(A14)}$$

For small $Q(QR_E < 1)$ the second and third terms are negligible and $S_{chain}(Q,t)$ describes the center of mass diffusion of the chain

$$S_{self}(Q,t)=S_{chain}(Q,t)=\exp\left(-D_RQ^2t\right) \qquad \text{(A15)}$$

For $Q\,R_E>1$ and $t<\tau_R$ the internal relaxations dominate. For $t=0$ we have $S_{chain}(Q,t) = S_{chain}(Q)$ i.e. the structure factor corresponds to a snapshot of the chain structure

$$S_{chain}(Q)=\frac{1}{N}\sum_{n,m}\exp\left(-\frac{1}{6}Q^2|n-m|\ell^2\right) \qquad \text{(A16)}$$

Replacing the summation by integrals and observing the relation $R_g^2 = \frac{1}{6} N\ell^2$ for the radius of gyration Eq.[A16] immediately leads to the well known Debye function.

$$S_{chain}(Q) = N f_{Debye}\left(Q^2 R_g^2\right)$$
$$f_{Debye}(x) = \frac{2}{x^2}\left(e^{-x} - 1 + x\right) \qquad \text{(A17)}$$

The shape of the Debye function corresponds to the upper most curve in Figure A2.

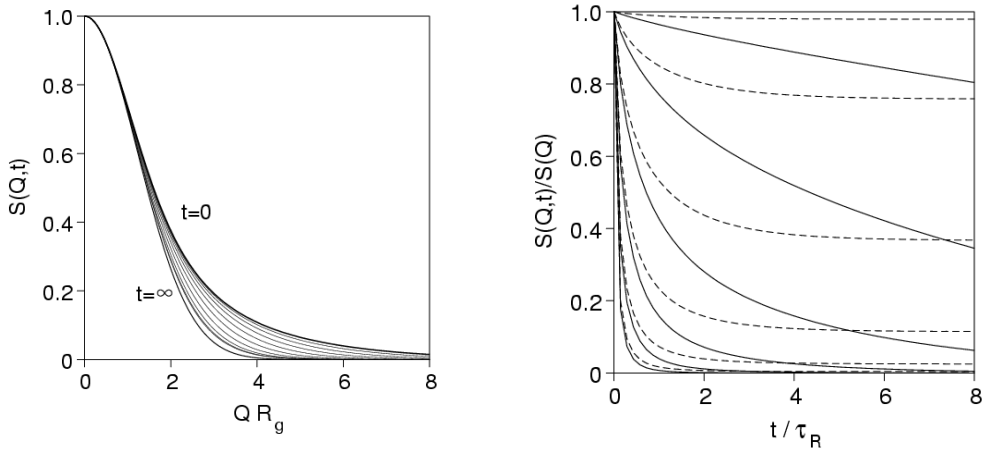


Fig. A2: *Development of S_{chain} (Q,t) for different times (left) and the normalized relaxation function $S_{chain}(Q,t)/S(Q)$ (right) for QR_g = 1, 2, ... 6. The dashed lines contain only the intra-chain relaxation, whereas the solid lines include the center-of-mass diffusion. Note that for short chains, respectively for small Q the translational diffusion dominates the observed dynamics.*

Important scaling properties are revealed by an approximate computation of the high Q behaviour of $S_{chain}(Q,t)$ (Eq.[A.14]). Replacing sums by integrals and performing some simplifications we get

$$S_{chain}(Q,t) = \frac{12}{Q^2\ell^2}\int_0^\infty du \exp\left\{-u - \left(\Omega_R t\right)^{1/2}\ h\left(u\left(\Omega_R t\right)^{-1/2}\right)\right\}$$
$$h(y) = \frac{2}{\pi}\int_0^\infty dx\ \frac{\cos(xy)}{x^2}\left(1 - \exp\left(-x^2\right)\right) \qquad \text{(A18)}$$

Note that this equation only depends on one variable: the Rouse variable.

$$\left(\Omega_R t\right)^{1/2} = \frac{Q^2}{6}\sqrt{\frac{3k_B T\,\ell^2 t}{\zeta_0}} = \frac{Q^2\ell^2}{6}\sqrt{Wt} \qquad \text{(A19)}$$

The Rouse model does not contain an explicit length scale. Therefore for different momentum transfers the dynamic structure factors are predicted to collapse to a single master curve if they are represented as a function of the Rouse variable.

Further note, that for $t = 0$ Eq.[A.18] does not resemble the Debye function but yields its high Q limiting behaviour $\approx Q^{-2}$. The approximation is only valid for $QR_g \gg 1$. In that regime the form of Ω_R immediately reveals that the intrachain relaxation increases $\approx Q^4$ in contrast to normal diffusion $\approx Q^2$. Finally, Figure A.2 illustrates the time development of the structure factor.

7 References

[1] For a recent review see: D. Richter, M. Monkenbusch, A. Arbe, J. Colmenero, Adv. Polymer Sci. 174, Springer, Berlin, Heidelberg, New York (2005)
[2] M. Doi, S.F. Edwards, J. Chem. Soc. Farad. Trans. 274, 1789; 274, 1802; 275, 38 (1978)
[3] P.G. de Gennes, J. Chem. Phys. 55, 572 (1971)
[4] Z. Bu, R. Biehl, M. Monkenbusch, D. Richter, J.E. Callaway, PNAS 102, 17646 (2005)
[5] D. S. Pearson, L. J. Fetters, W. W. Graessley, G. V. Strate, E. v. Meerwall, Macromolecules, 21, 711 (1994)
[6] T. Lodge, Physi. Rev.Lett., 83, 3218 (1999)
[7] P.R. Rouse, J. Chem Phys. 21, 1272 (1953)
[8] F. Mezei (Ed.), Neutron spin echo lecture notes in physics 128, Springer, Berlin, Heidelberg, New York (2005)
[9] D. Richter, B. Ewen, B. Farago, T. Wagner, Phys. Rev. Lett., 62, 2140 (1989)
[10] H. Montes, M. Monkenbusch, L. Willner, S. Rathgeber, L.J. Fetters, D. Richter, J. Chem. Phys. 110, 10188 (1999)
[11] K. Kremer, G. S. Crest, J. Chem. Phys. 92, 5057 (1990)
[12] M. Doi, S.F. Edwards, The theory of polymer dynamics, Clarendon Oxford (1986)
[13] P.G. de Gennes, Scaling concepts in polymer physics, Cornell University Press, Ithaca (1979)
[14] N. Fatkullin, R. Kimmrich, Phys. Rev. E 59, 3273
[15] A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, B. Farago, G. Kali, Phys. Rev. Lett. 90, 058302 (2003)
[16] P.G. de Gennes, J. Phys. (Paris) 42, 735 (1981)
[17] P. Schleger, B. Farago, A. Kollmar, C. Lartigue, D. Richter, Phys. Rev. Lett. 81, 124 (1998)
[18] N. Clarke and T. C. B. McLeish, Macromolecules 26 5264 (1993)
[19] A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, A. E. Likhtman, T. C. B. McLeish and B. Farago, Phys. Rev. Lett. 88, 058301 (2002)
[20] M. Zamponi, A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, A. E. Likhtman, G. Kali, B. Farago, Phy. Rev. Lett. 96, 238302 (2006)
[21] M.M. Tirion, Phys. Rev. Lett. 77, 1905 (1996)
[22] R. Biehl, B. Hoffmann, M. Monkenbusch, P Falus, S. Preost, R. Merkel, D. Richter, Phys. Rev. Lett. 101, 138102 (2008)
[23] N. Brown, Dynamic Light Scattering, Monographs of the Physics and Chemistry of Materials, Oxford Science Publications, Oxford 49 (1993)
[24] J. Garcia de la Torre, M.L. Huertas, B. Carrasco, Biophysics J. 78, 719 (2000)
[25] R. Inoue, R. Biehl, T. Rosenkranz, J. Fitter, M. Monkenbusch, A. Radulescu, B. Farago, D. Richter, Biophysical Journal 99, 2309 (2010)

E 4 Material Systems for Information Technologies [1]

M. Angst

Peter Grünberg Institut, Jülich Center for Neutron Science, and JARA-FIT

Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

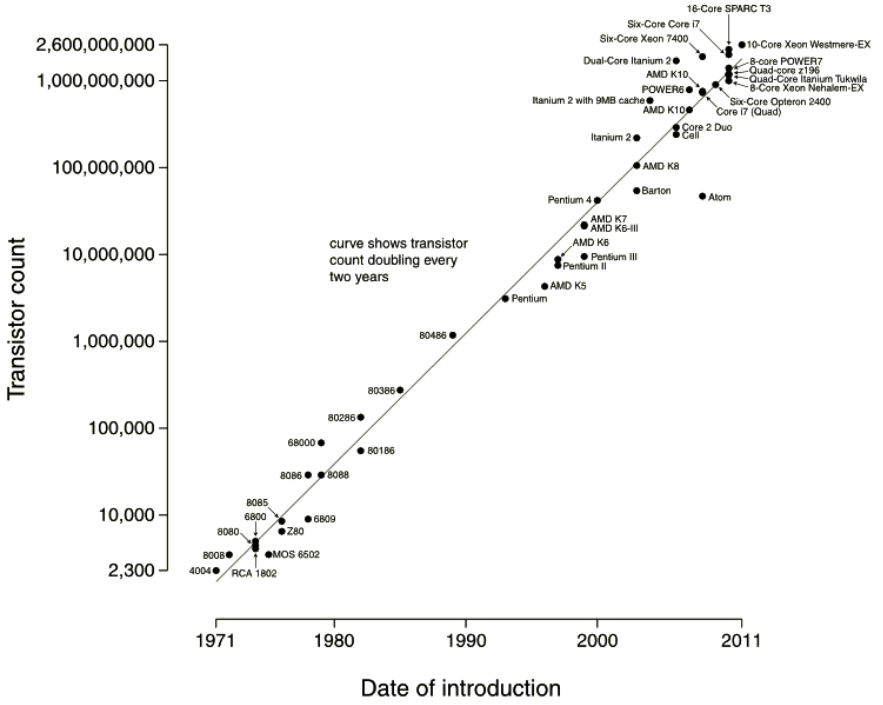


Fig. 1: *"Moore's law" describes the doubling of the number of transistors every two years (Image Source: Wikipedia).*

1 Introduction

Information technology (IT) increasingly permeates all areas of our society and life. It plays a key role in areas as varied as communication and organization, transport and logistics, energy and environmental technology, health and medical care, and last but not least entertainment. The long-term trends in IT are described by Moore's law: the number of transistors placed on integrated circuits doubles every two years (Fig. 1). Similar exponential dependencies also describe computing speed, memory capacity, and, on the downside, costs to build a chip factory and waste heat generated. Although Moore's law has now held for 40 years, it cannot be extrapolated to the future indefinitely. With increasing transistor density implying smaller transistors, about 2025 structures will extend over just a few atoms. This sets a fundamental limit for traditional silicon-based technology: for atomic length scales, quantum effects such as tunneling will hamper device reliability. Furthermore, the waste heat generated by operating these transistors will become unmanageable.

Continuing Moore's law over the next few decades will therefore require fundamentally new concepts and material systems beyond silicon. A large variety of concepts are under investigation, including e.g. quantum computing with quantum dots [1] or superconducting qubits [2], graphene-based transistors [3], redox-based resistive switching [4], spintronics with single-molecular magnets [5], or interfacing with biological systems [6]. Currently it is completely unclear, which of these concepts will succeed in the long run. As it is impossible to cover all of these concepts in more than a superficial way, I will instead focus in this lecture on two concepts, phase change materials and multiferroics, both having a high potential for being applied to vastly improved non-volatile memories already in the near future.

Desirable would be a "universal memory" combining cost benefits of DRAM (used for the main memory of computers), speed of SRAM (used for CPU cache), and non-volatility of flash memory (USB sticks). Furthermore, the ideal memory should have high information density, low power consumption, and a long lifetime (stable over many cycles). Currently the most

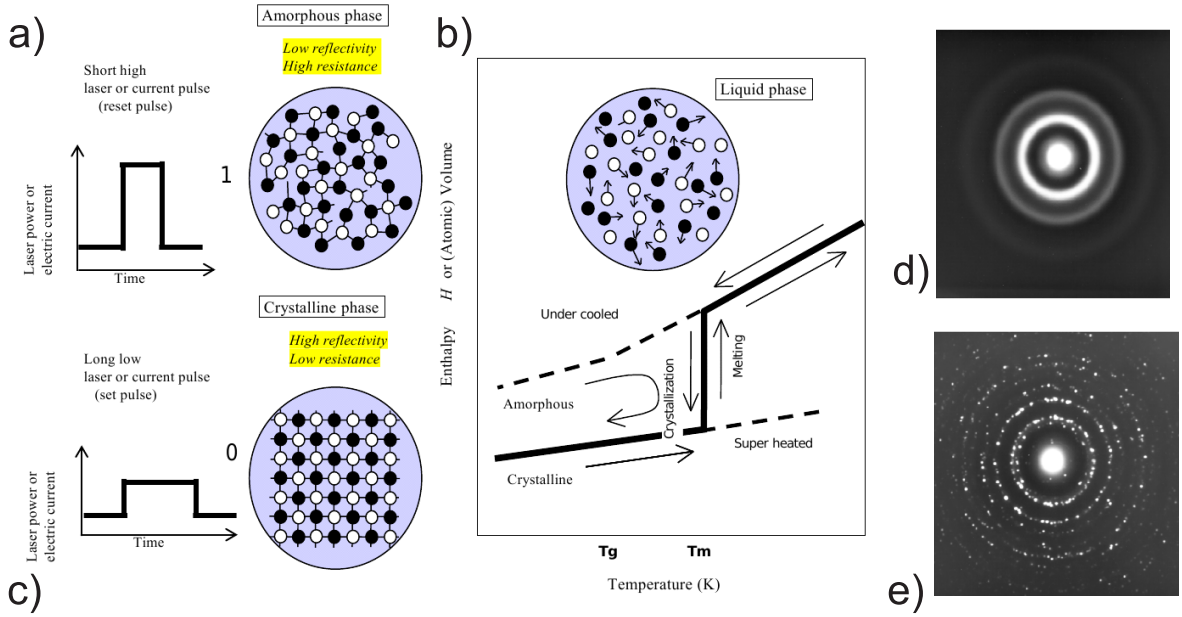


Fig. 2: *a-c) Rewritable optical data storage with PCM (after [7]). a) Short high laser pulse heats the PCM locally, melting it. b) Rapidly cooling quenches the liquid into an amorphous phase with different optical properties from the surrounding crystalline material (detectable by low-intensity laser pulse). c) A long lower-power laser pulse heats the material also to the liquid state, but the less rapid cooling results in crystallization. d/e) Phase contrast between amorphous phase (d) and laser-irradiation-induced polycrystalline state (e) in $Ge_2Sb_2Te_5$ by selected-area electron diffraction (after [8]).*

used non-volatile memory, flash, has high information densities, but a very slow and power-consuming erase operation, and it survives only about 10^5 cycles. Less often used are non-volatile memories based on ferroelectricity, which alleviates some of the drawbacks of flash, but has a low information density. Further, magnetic non-volatile memories (MRAM) are in use, the principle and main drawback of which is addressed in Sec. 3.1.

In the following, first phase change materials (PCM) will be discussed as a route towards an "universal memory" (Sec. 2). After describing the operating principle and the extensive present use of PCM for rewritable optical disks, the main focus is on the unusual structure and bonding underlying the functionality. Next, multiferroics will be considered (Sec. 3). After introducing the concept and the operation principle, the main focus will be on the problem of the incompatibility between traditional ferroelectricity and magnetism, and various mechanisms providing a way around it. The lecture will conclude with a comparison of the application prospects of PCM and multiferroics and a brief discussion of the relevance of scattering methods in elaborating these materials.

2 Phase Change Materials (PCM)

2.1 Rewritable disks and non-volatile memories with PCM

Phase-change materials (PCM) are materials with metastable amorphous and crystalline phases, and fast transformations between them. They are already used in IT, in rewritable optical data

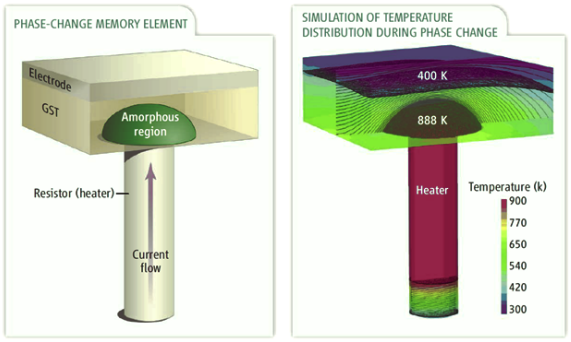


Fig. 3: *Concept of nonvolatile phase-change memory cell. Sketch of the cell and resistive heating to write/erase (left) and simulation of temperature-distribution during phase change. After [11].*

storage media such as newer generations of rewritable DVDs (digital versatile disks) and Blu-ray disks. For such media, the bits are encoded as sections of the PCM in either a crystalline (0) or an amorphous (1) state, both of which have a thermal stability of several decades at room-temperature. For PCM used in optical data storage, considerable differences in the atomic arrangement between amorphous and crystalline phases results in a significant optical contrast. The state of the bit can then be read out by a very low-power laser pulse.

The operating principle of PCM is illustrated in Fig. 2a-c). To write a bit, a short high laser pulse is applied, which heats the material above its melting temperature. Because the melting temperature T_m is far above ambient temperature, the PCM cools very rapidly below the glass forming temperature T_g after the pulse, at rates $\sim 10^9\,\mathrm{Ks}^{-1}$. This leads to quenching of the liquid state into a disordered, amorphous phase. The reverse operation also utilizes heating by a laser pulse, but one that is of lower power over a longer duration, heating the PCM to $T_g \ll T < T_m$. The long pulse duration giving sufficient time for crystallization of the PCM. The phase transformation from amorphous to crystalline by laser irradiation is directly verifiable e.g. by selected-area electron diffraction, see Fig. 2d/e) [8].

For the application several requirements need to be fulfilled: In addition to the long thermal stability of both states at room temperature and to the large optical contrast, suitable materials must be glass formers, but marginal ones, so that both phases can be reached easily. This means typically reduced glass temperatures T_g/T_m between 0.5 and 0.55 [9]: For smaller T_g/T_m it becomes difficult to reliably quench the PCM rapidly enough to reach the amorphous state. For higher T_g/T_m, on the other hand, the crystallization rate is too slow for the material to be of practical use. Additional requirements include a good reversibility of the transition, i.e. a large number of cycles without degradation, and a high chemical stability [7]. Successful examples of PCM include Te-based multicomponent alloys along the GeTe-Sb_2Te_3 quasi-binary tie-line, such as $Ge_2Sb_2Te_5$ [10].

Optical data storage is not the only possible IT application of PCM. Indeed, PCM show also great promise for future non-volatile memory-cells [11]. The operating principle in this case is very similar to the one for optical disks, with the optical contrast between the phases replaced by different (low-current) resistivities and the laser pulse heating replaced by current pulse heating (Fig. 2a-c). The current pulse is applied to a resistive heater placed next to the PCM-region of the cell, as indicated in Fig. 3.

Because strong resistive contrasts are much easier to reach than significant optical contrasts, there could be many more candidate PCMs for non-volatile memory cells than for rewritable optical storage media. In this context I would like to note that historically, resistive switching was how PCM were discovered several decades ago [12], while their use for optical disks was noted only much later.

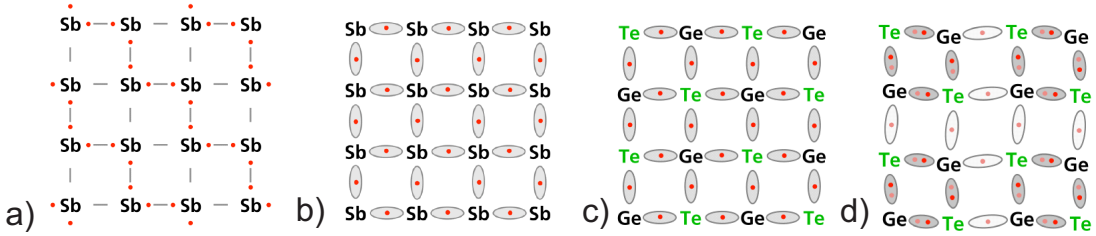


Fig. 4: *Resonant bonding and distortions in Sb and GeTe (2D-projection). a) With covalent p-bonding in Sb, only half of the bonds are occupied. b) Sb resonant bonding with delocalization of the electrons in the bonds. c) Resonant bonding in GeTe. d) as c), but with additional slight distortion. After [14].*

However, the usability for non-volatile memories involves some requirements beyond those for the optical disks. First, the speed of the write and also the erase operations is even more crucial. This implies for example very fast crystallization dynamics, as this is the speed-limiting factor. Crystallization of the order of $100\,\mathrm{ns}$ has been achieved so far [11], which is already two orders of magnitudes better than a cycle in a flash memory. A long lifetime (large number of cycles) is another requirement, which is already orders of magnitude better. Ion migration under the influence of the electric field during switching is the major factor preventing even longer lifetimes [13].
Crucial for operation is also a minimization of waste heat, i.e. the switching should not require unduly high power. Resistivity in the crystalline phase are typically Ohmic and intermediate, which ideally facilitates resistive heating. However, in the amorphous phase the resistivity is extremely high, preventing enough current to flow for the necessary heat to be delivered. Fortunately, at moderate voltages ($< 1\,\mathrm{V}$ for typical dimensions, threshold switching occurs in the amorphous phase [9]: a fast electronic transition into state with much lower resistance. The subsequent current flow then generates the required heat.

2.2 Structure and Bonding

Because the phase change is a structural change there must be a strong relationship between the PCM functionality and the structures of both the crystalline and the amorphous phases. Thus, the structure is examined here, starting with the crystalline phase. For e.g. $Ge_2Sb_2Te_5$ the thermodynamically stable crystal structure is trigonal, but the phase involved in phase change is exclusively a metastable cubic phase (as indicated by diffraction such as the pattern shown in Fig. 2d). This cubic phase corresponds to a slightly distorted NaCl-structure, with the Cl-positions occupied by Te and the Na-positions randomly by Ge, Sb, and vacancies. Thus all the atoms have octahedral (6-fold) coordination. This octahedral coordination is typical for PCM in the crystalline state, but in general highly unusual, as it violates the so-called "$8 - N$ rule" that states that the coordination of an atom in a covalent bonding environment should be $8 - N$ when N is the number of valence electrons. Thus, Ge should be 4-fold coordinated (as is the case for elemental Ge), Sb 3-fold and Te 2-fold.
Consider Sb, also a PCM, in the ideal NaCl-structure. Sb has a $5s^2 5p^3$ valence shell configuration. The $90°$ bond angles for the NaCl structure suggest the absence of $s - p$ hybridization, i.e. only the p-orbitals take part in bonding. With normal covalent bonding only half of the possible bonds are occupied, e.g. those shown in Fig. 4a), or completely equivalently all those

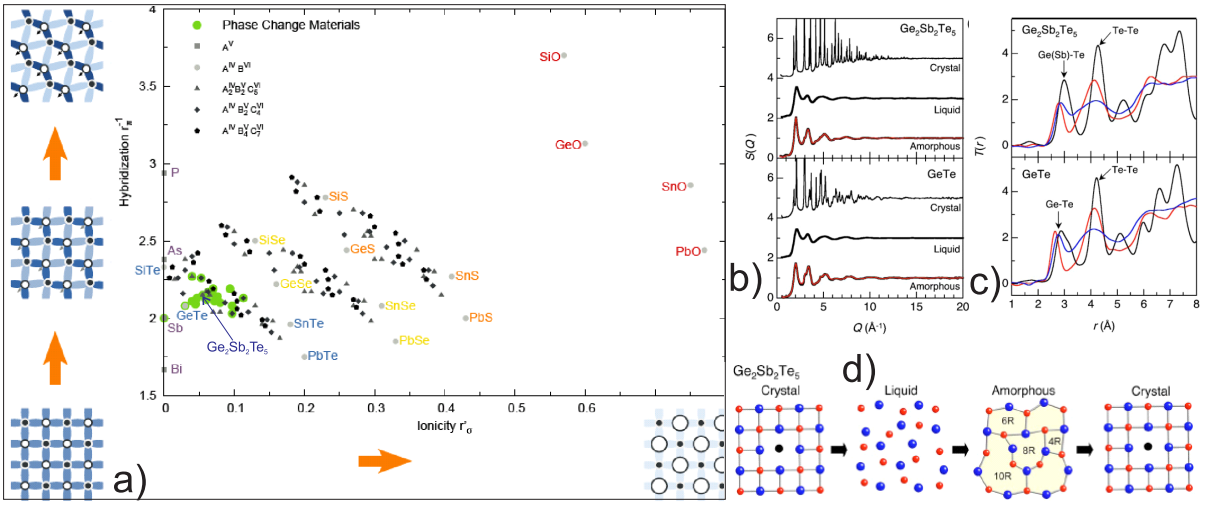


Fig. 5: *a) Treasure map for PCM (after [16]): The x-axis is the ionicity, the y-axis the tendency towards hybridization. PCM are found in a small region of the map that is prone to resonant bonding. b)-d) Local structure of $Ge_2Sb_2Te_5$ and GeTe phases (after [17]). b) total structure factors. c) pair correlation functions for crystal (black), liquid (blue), and amorphous (red) phases. d) Schematic representation of possible ring-size transformation in recording and erasing processes ($Ge_2Sb_2Te_5$).*

not occupied in Fig. 4a). In this situation, a superposition of the two bonding pattern occurs, with highly delocalized electrons (Fig. 4b). This is called resonant bonding [15] and occurs more prominently in benzene. For PCM like GeTe, the picture is analogous (Fig. 4c). Alternatively to resonant bonding, a Peierls distortion can occur, involving alternating short and long bonds, deviation of the angles from $90°$, and the opening of a gap. For large distortions, resonant bonding is no longer possible. But for small distortions (sketched for GeTe in Fig. 4d) as typically occur in the crystalline phase of PCM, the electron localization into half of the bonds is incomplete, and resonant bonding is only weakened. The resonant bonding, even weakened, leads to a high electronic polarizability and low resistivity and is thus the origin of the property contrast in PCM as the transition to the amorphous phase involves a much more distorted state (see below) destroying resonant bonding [14].

To pinpoint the occurrence of resonant bonding, Lencer and coworkers [16] considered a 2D map of materials (Fig. 5a). One coordinate axis is the degree of ionicity of the bonding, which can be parametrized as $r'_\sigma = r_p^{\mathrm{A}} - r_p^{\mathrm{B}}$, where r_p^{A} and r_p^{B} denote the averaged valence radius of the p-Orbital for cations and anions, respectively. The other axis gives the tendency towards $s-p$ hybridization, parametrized as $r_\pi^{-1} = [(r_p^{\mathrm{A}} - r_s^{\mathrm{A}}) + (r_\pi^{-1} = [(r_p^{\mathrm{B}} - r_s^{\mathrm{B}})]^{-1}$, with s denoting the s-orbitals. For high values of r_π^{-1} $s-p$ hybridization becomes inreasingly favorable, resulting in a deviation from octahedral connectivity and coordination number following the $8-N$ rule (see sketches at the left). This excludes resonant bonding. Similarly, increasing ionicity localizes the electrons at the anions (see bottom sketches), also opposing resonant bonding. We thus expect PCM only in the lower left corner of Fig. 5a), which is indeed the case (green dots). This confirms the high relevance of resonant bonding and also provides a tool in the search of further PCM. Whereas the PCM-region is at minimal ionicity, it includes a small but non-zero amount of hybridization tendency. Therefore, while resonant bonding must prevail, it also has to be imperfect, with some propensity towards hybridization.

To understand the strong property contrast between crystalline and amorphous phases despite

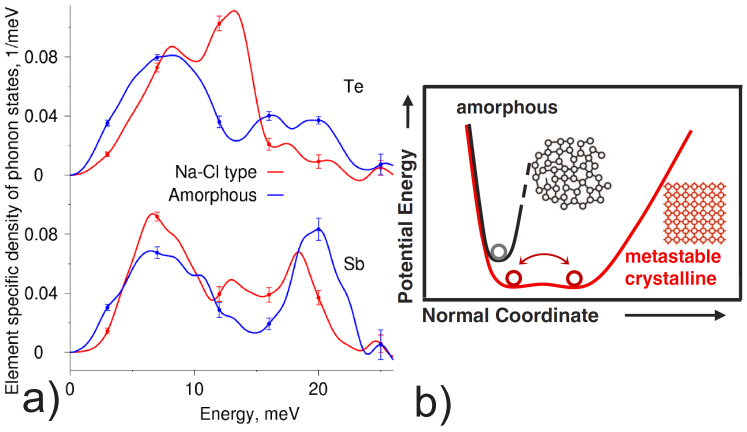


Fig. 6: *a) Elastic hardening and vibrational softening of optical modes: Crystalline (red) and amorphous (blue) $GeSb_2Te_4$ element-specific phonon density-of-states from nuclear-inelastic scattering at* 25 K. *b) Schematic interaction potential for the amorphous (black) and crystalline (red) state of a PCM. After [19].*

of the fast transformations between them, it is also necessary to elucidate the local structure in the amorphous phase. This can be done by total scattering or pair-distribution-function analysis (see lecture D5) or by x-ray absorption spectroscopies (see lecture F4). An example of the application of the former to $Ge_2Sb_2Te_5$ and GeTe is shown in Fig. 5b)-d) [17]: the radial total structure factor of crystal, liquid, and amorphous, phases obtained by x-ray scattering (panel b) by fouriertransform yields the radial pair-correlation functions $T(r)$ (probability to find a second atom at distance r), see panel c). Compared to the crystalline phase (black), $T(r)$ of the amorphous phase (red) is washed out particularly at large r, as expected. However, unusually, there are also pronounced shifts of the peaks indicating the nearest-neighbor distances from Ge(Sb) to Te. Performing Reverse Monte Carlo on the data, Kohara and coworkers obtained the short and intermediate range structure in all $Ge_2Sb_2Te_5$ phases (Fig. 5d). The average coordination numbers for Ge (3.7), Sb (3.0), and Te (2.7) are much closer to the expectation from the $8 - N$ rule than for the crystalline phase, suggesting essentially covalent bonding with $s - p$ hybridization. Despite of these coordination numbers, the bond angle distribution (not shown) peaks at $90°$, indicating a preferentially octahedral environment of the atoms, even if highly defective and distorted. The structure can be described by statistics on ring sizes, with each ring showing pronounced cation-anion alternation (as in the crystalline cubic phase). To a first approximation, the amorphous structure can thus be treated as a highly distorted copy of the cubic phase, but the strong distortion removes the resonant character of the bonding, resulting in the property contrast. Molecular dynamics simulations [18], simulating a whole phase-change cycle, indicate that there is a very high density of 4-rings in the amorphous phase, involving more than 50% of the atoms. These serve as crystallization centers, facilitating the fast transformation to the cubic phase, for which little medium-ranged diffusion is necessary.

2.3 Vibrational properties

Particularly for applications in non-volatile memories, a low thermal conductivity is also important, because it decreases the power necessary to thermally cycle a memory cell. Low thermal conductivities in amorphous phases is typical, but in contrast to most materials, PCM have thermal conductivities in their crystalline phase that is not much higher than in the amorphous phase. This points to unique vibrational properties of the crystalline phases of PCM. These can be assessed e.g. by nuclear inelastic scattering (see lecture D9) yielding the element-specific phonon density-of-state, shown for both phases of the PCM $GeSb_2Te_4$ in Fig. 6a) [19]. Surprisingly, crystallization involves both a hardening of the low-energy acoustic phonons (and

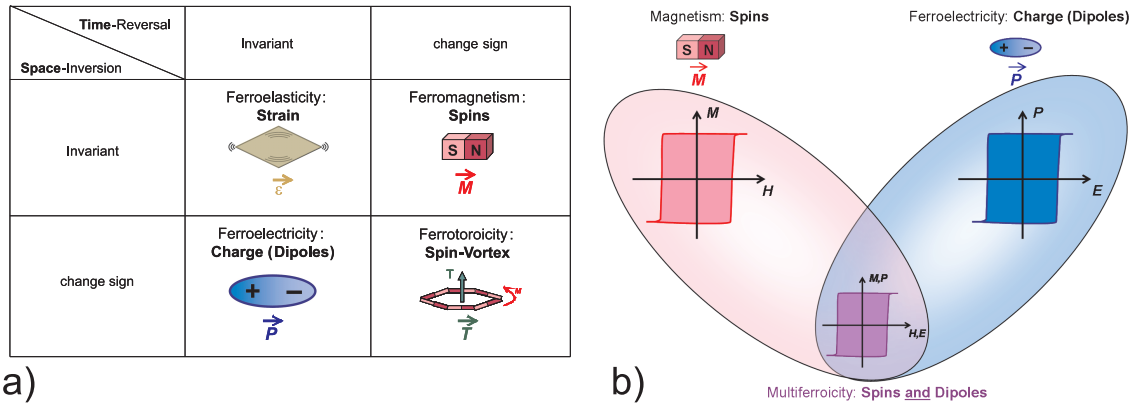


Fig. 7: *a) Classification of "ferroic" orders according to symmetry. b) "Ferroic" also requires coupling of the order parameter to an appropriate field, e.g.* $M(H)$ *or* $P(E)$*, without which the order is termed "antiferroic". "Multiferroics" as a term was coined to label materials combining two or more ferroic orders. In particular, magnetoelectric Multiferroics combine ferromagnetic (FM) and ferroelectric (FE) orders, and the both order parameters may couple to both magnetic and electric fields.*

thus a larger bulk modulus) and a softening of the high-energy optical phonons. The hardening of acoustic phonons corresponds to an increase of the acoustic-mode Debye temperature θ, but only a moderate ($\sim 20\%$) one. The unusual simultaneous softening of the optical phonon modes is an effect of the resonant bonding in the crystalline phase [19]. Because some of the optical phonons are closely related to the tendency towards a Peierls-distortion, this softening suggests a small barrier between the undistorted cubic structure and a Peierls-distorted structure. The interaction potential in the crystalline phase looks thus qualitatively as sketched in Fig. 6b) in red: it is weaker and much more anharmonic than the corresponding potential in the amorphous phase. Consequences of this interaction potential include significant static distortions (observed in form of large displacement parameters) and large thermal expansion (as also observed). Furthermore, the anharmonicity, quantified as the high-T limit of the Grüneisen-Parameter γ, in inversely proportional to the thermal conductivity $\kappa \propto \theta^3\gamma^{-2}$, and thus contributes to the anomalously low thermal conductivity in the crystalline phase. Because the unusual interaction potential arises from the combination of resonant bonding and static distortions, the PCM-region in the treasure map (Fig. 5a) with its low, but not too low propensity towards hybridization, is also the region of solids with low Debye temperatures, anharmonic interactions, and low thermal conductivities, i.e. PCM fulfill this requirement automatically.

3 Multiferroics

3.1 Multiferroics for non-volatile memories

The term "multiferroic" (MF) was coined [20] to describe materials simultaneously exhibiting at least two kinds of ferroic orders. A ferroic order is described as a spontaneous order that can exhibit domains (where the order parameter has different sign or direction), and whose order parameter couples to a suitable external field. For example, in a ferromagnet (FM) the Magnetization (order parameter) can form domains, and it can be switched by (couples to) a magnetic field, represented by a hysteresis loop as sketched in red in Fig. 7b). Ferroic orders can be

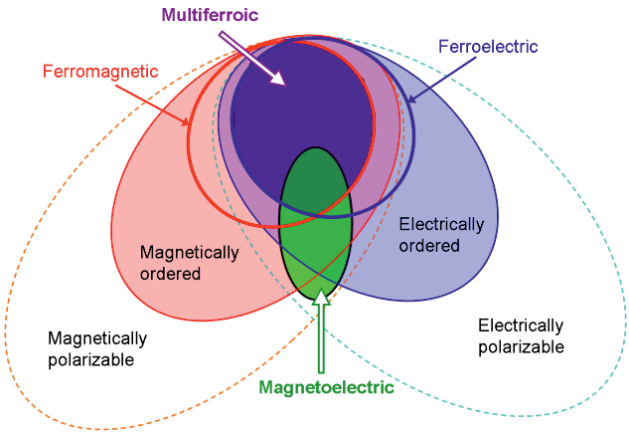


Fig. 8: *Often the term "multiferroic" is used in a wider sense to encompass simultaneously magnetically and electrically ordered materials, e.g. antiferroelectric antiferromagnets. A magnetoelectric coupling can exist within the even wider class of both magnetically and electrically polarizable materials.*

classified according to the transformation-properties of the order parameter upon time-reversal and space-inversion (Fig. 7a). For example, ferroelastic order is invariant under both operations, whereas the order parameter of ferrotoroidicity (the alignment of toroidal moments arising from a ring-like arrangement of spins, see [21] and Fig. 7a) changes sign upon either operation. A ferroelectric (FE) that is also a ferromagnet (magnetoelectric MF) breaks both time-reversal and spatial-inversion symmetries, similar as a ferrotoroidic. The latter are also interesting for magnetoelectric applications, as they intrinsically exhibit a magnetoelectric effect. Research on ferrotoroids is relatively new, and this fascinating topic cannot be covered in detail here, see [21] for a comprehensive review. Despite being relevant for applications, we also omit ferroelastics from further consideration, focusing entirely on magnetoelectrics from now on, writing just multiferroic or MF for magnetoelectric multiferroic.

The combination of FM and FE into a multiferroic is sketched in Fig. 7b). By definition then, such a material has a spontaneous magnetization that is switchable by a magnetic field and a spontaneous electric polarization that is switchable by an electric field. With a sufficient electromagnetic coupling (which is not guaranteed), it is possible that the spontaneous magnetization can also be switched by an electric field, or the polarization by a magnetic field. However, multiferroicity does *not* imply that there is a magnetoelectric effect , and conversely magnetoelectric effects can occur in the much wider class of materials that are simultaneously magnetically and electrically polarizable (Fig. 8). In practice, magnetic and electric ordering tendencies will make a strong magnetoelectric coupling much more likely, however, as indicated above. Since actual FM FE are more rare than antiferromagnetic (AFM) FE, most practitioners have come to use the term "multiferroic" in a wider sense, i.e. designating any magnetically ordered FE (or even including anti-ferroelectric magnetic materials as well). We will follow this convention through the rest of this section.

As mentioned in Sec. 1, multiferroics may also be used for non-volatile memories. Consider the MRAM or magnetic random access memory, a non-volatile memory already in use. Schematically, the already existing MRAM cells look similar to the device depicted in Fig. 9: Information is encoded in the relative direction of the magnetization of two layers (parallel: 0, antiparallel: 1), and it is read out with the giant magnetoresistance (GMR) effect, discovered in 1988 (P. Grünberg, Forschungszentrum Jülich, and A. Fert, U Paris-Sud 11, Physics Nobel Price 2007). The main problem of the MRAM is the write operation requiring remagnetizing a layer with significant currents, making it slow and in particular highly power consumptive. If a layer of an insulating material that converts a voltage to a magnetization could be attached, writing the device could occur with the simple and non-dissipative application of a voltage. The resulting

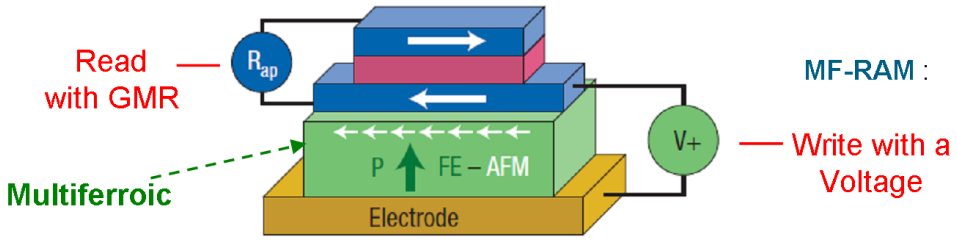


Fig. 9: *Concept of multiferroic memory cell [22]. As in the already existing magnetic RAM (MRAM), the information is encoded in the parallel or antiparallel alignment of the magetization of two layers, read out with the giant magneto-resistance effect. The addition of a layer of magnetoelectrically active (e.g. multiferroic) material allows to write the information by applying a voltage, overcoming the main drawback of MRAM cells.*

cell might be called "multiferroic" RAM or MF-RAM. The principal challenge in realizing the MF-RAM concept is to find a suitable material, for which a magnetoelectric coupling as described above occurs at room temperature and it is strong enough. Luckily, in order to switch the magnetization it is not necessary to have a spontaneous magnetization in the magnetoelectrically active (green in Fig. 9) layer. This is because of a directional magnetic coupling at the interface between adjacent FM and AFM layers that is known as exchange bias (reviewed in [23]). Due to the directionality the magnetization direction in the ferromagnetic layer may follow the direction of the AFM order parameter in the AFM layer. Bi-stable switching of magnetization by an electric field has indeed been demonstrated using exchange-bias and an AFM FE (see Sec. 3.5). However, although AFM instead of FE works, identifying suitable materials is still a challenge.

3.2 Incompatibility between Ferroelectricity and Magnetism and ways around

There are a lot of FM materials as well as a lot of FE ones. However, there are surprisingly few that are both, and understanding why is the first issue to be addressed [24]. One reason for the scarcity of multiferroics is simply symmetry: due to the need of breaking both time-reversal and spatial-inversion symmetries, out of the 122 magnetic point groups there are only 13 that allow for both spontaneous magnetization and electric polarization to occur. Yet, many materials occur in one of these 13 point groups without being multiferroic. A simple reason for the rarity of multiferroics in the strict sense (FM FE) is that ferromagnets tend to be metals whereas sustaining an electric polarization requires a material to be insulating. More interesting is to examine the wider class of magnetic FE, examples of which still turn out to be very rare.
To address this latter question, we focus on perovskite-type ternary oxides ABO_3 with A and B transition metals, which contain both numerous (hundreds) FE and magnetic materials. At high temperatures perovskites have an often slightly distorted, but centrosymmetric, cubic structure. The typical, and "traditional", mechanism of ferroelectricity in these materials is an off-center displacement of the B-site ion (for example Ti in $BaTiO_3$), sketched in Fig. 10a). With a $4+$ valence from electron counting, Ti has an empty d shell. In fact, most FE perovskites have a B ion with empty d shell. Since magnetism requires unpaired d electrons, this automatically prevents multiferroicity in these cases.
Since electrostatic interactions (Coulomb-repulsion between electron clouds on adjacent ions)

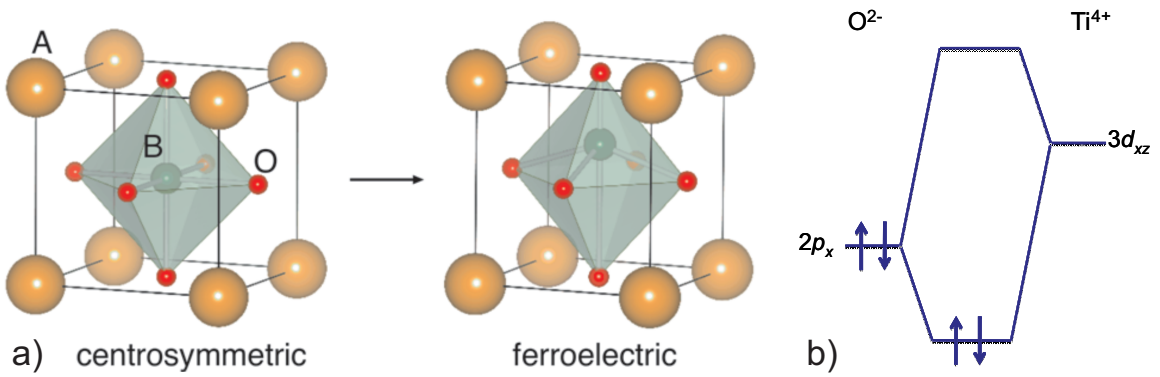


Fig. 10: *"Traditional mechanism of ferroelectricity": a) In perovskites ABO_3 such as $BaTiO_3$, an off-center displacement of the B ion leads to ferroelectricity. b) B-O Hybridisation energy level scheme. If B has an empty d-shell, the shared electrons occupy exactly the bonding state (arrows). Additional electrons from the B d-shells would go into anti-bonding orbitals, lowering any energy gain from hybridisation.*

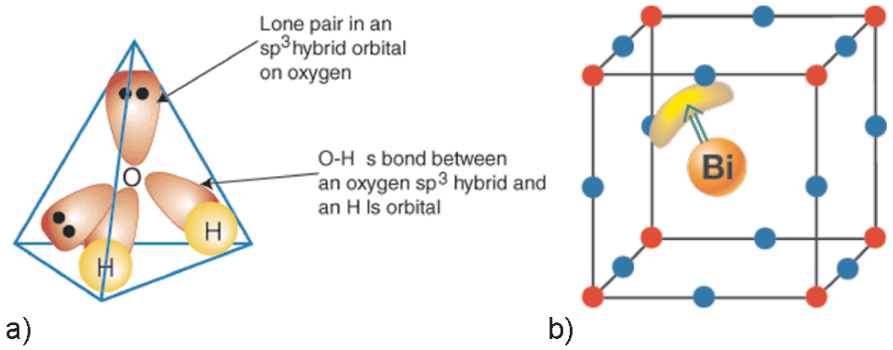


Fig. 11: *a) Oxygen lone pairs in a water molecule. b) Bismuth lone pair in $BiCrO_3$ [27]. The lone pair, if localized, breaks inversion symmetry and can lead to ferroelectricity.*

favor the centrosymmetric arrangement of ions, the B ion displacement must be due to other bonding considerations [24, 25]. Indeed, energy may be gained by formation of strong covalent bonds with one (or three) of the surrounding oxygen ions, with the energy gain from hybridization increased by decreasing the distance. Hybridization of the corresponding Op and and Bd orbitals leads to a bonding and an antibonding "molecular" orbital, as depicted in Fig. 10b). If the d shell of the B ion is empty, the two electrons from the O^{2-} ion (arrows) exactly fill the bonding orbital, maximizing the energy gain from hybridization. In contrast, any d electrons from B would have to go into the anti-bonding orbital, lowering the energy gained.

While the above consideration makes plausible the strong preference for Bd^0 configurations for this "traditional mechanism of ferroelectricity", I should remark that it is no absolute theorem and additional factors likely participate, but the detailed discussion we have to omit here [25]. In any case, the connection between the "traditional mechanism of FE" in perovskites and the "d^0th-ness" of the B ion is established empirically.

The challenge, then, is to combine magnetism and FE despite the above-mentioned incompatibility. The conceptually simplest approach is to fabricate multi-phase materials or thin film hetero-structures, using $BaTiO_3$ or another traditional FE and a magnetic and magnetostrictive material such as $CoFe_2O_4$. Before briefly discussing this approach in subsection 3.5, we will focus on single-phase multiferroics (in the wider sense of magnetically and electrically ordered materials). The focus here will be on the experimental side; for a review of different approaches focusing on theory see [26].

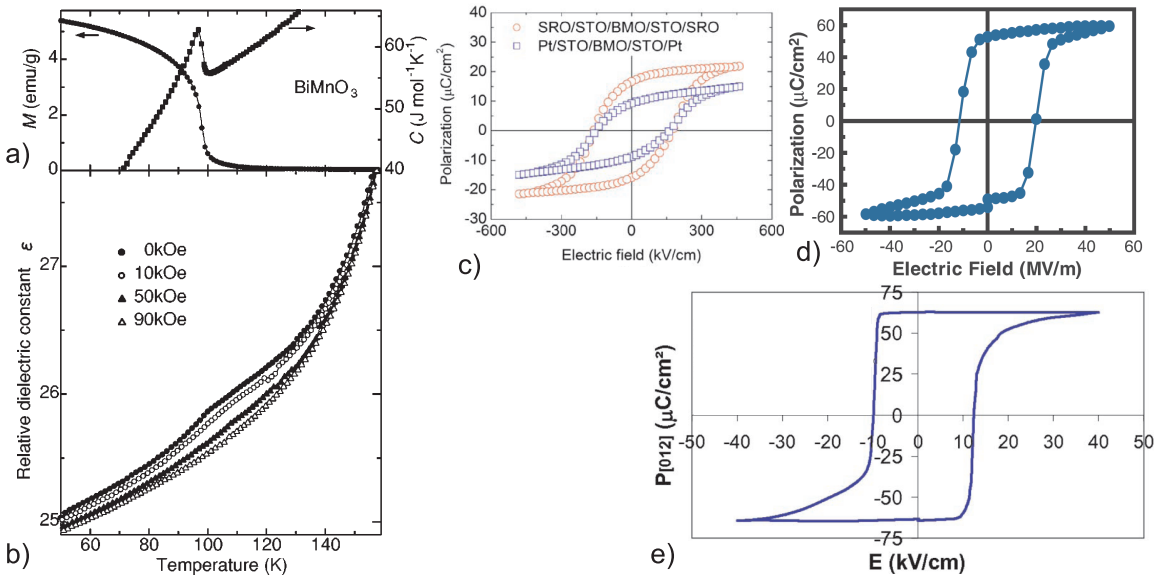


Fig. 12: *a) $BiMnO_3$ specific heat and low-field magnetization [30]. b) $BiMnO_3$ dielectric constants in different magnetic fields [30]. c) $BiMnO_3$ Polarization [31]. d/e) $BiFeO_3$ thin film [32]/single crystal [33] polarization.*

The straight-forward way to achieve both magnetism and FE in a single structure is for the former to to originate from one subunit or ion, and the latter from another. For perovskites, one can for example on the B site partially substitute a ferroelectrically active (d^0) ion with a magnetic ion. The general main problem with such an approach is that due to the different ions (or subunits) involved in magnetism and FE, the magnetoelectric coupling tends to be very low in most cases [28].

As a concrete example of the "independent subsystems" approach we consider FE due to lone pairs at the A site in perovskites with magnetic B ions [29]. Lone pairs are electrons not used in chemical bonds, occurring for example in water molecules as shown in Fig. 11a). Lone pairs are highly polarizable, contributing e.g. to the polarizability of water. In perovskites, A site Bi^{3+} or Pb^{2+} have $6s$ electrons not participating in bonds and thus yielding lone pairs, see Fig. 11b). If these lone pairs are localized and ordered in one direction, inversion symmetry is broken and an electric polarization results. In $PbTiO_3$ this lone-pair mechanism helps the traditional mechanism stabilizing the FE.

In $BiMnO_3$, the B site Mn^{3+} ions lead to FM below $100\,\mathrm{K}$ (see Fig. 12a). The Bi lone pairs alone are sufficient to also stabilize FE (below $800\,\mathrm{K}$), as indicated by polarization hysteresis loops (Fig. 12c). The occurrence of both *ferro*electricity and *ferro*magnetism make this an exceptional example of a real multiferroic in the strictest definition [34]. Despite of this, however, the magnetoelectric coupling is very weak, as indicated by an only very small (and almost magnetic-field-independent) feature in the dielectric constant at the magnetic ordering temperature (Fig. 12b) [30]. The smallness of the dielectric feature at the magnetic transition is consistent with the expectation for an "independent subsystem" multiferroic.

In $BiFeO_3$ [35], the Fe spins order antiferromagnetically below $643\,\mathrm{K}$, in a complex spin structure based on G-type antiferromagnetism (i.e. with each Fe ion surrounded by six antiparallel nearest neighbors) [36]. Bi lone pairs again lead to FE, below $1100\,\mathrm{K}$, with polarization hysteresis loops both on thin films and more recently on high-quality single crystals (panels d,e) showing the intrinsic nature of a high ($> 60\,\mu\mathrm{C/cm^2}$, comparable with $BaTiO_3$) polarization at

room temperature [32, 33]. Room-temperature MF would make $BiFeO_3$ a very good prospect for applications if a significant magnetoelectric coupling were present as well. In analogy with $BiMnO_3$, we would expect only very small coupling. However, sizeable coupling and control of magnetism by electric fields has been demonstrated experimentally. The origin of this is connected with a spiral part of the magnetic structure and corresponding couplings in the class of "spiral multiferroics". I will therefore return to this issue later.

There are many more examples of "independent subsystem" multiferroics, which typically have FE transitions at much higher temperature than the magnetic transitions, and can have sizeable electric polarization [28]. The main challenge in this group of materials is to achieve (and understand) a significant magnetoelectric coupling.

3.3 Ferroelectricity from Spin-Spirals

In the next class of multiferroics to be considered [39, 40], the situation is reversed. These materials have FE transitions at the same or, more typically, lower temperatures as the magnetic transitions, typically show large magnetoelectric couplings as observed by magnetic-field effects on dielectric constants and polarization, but conversely have much smaller electric polarizations.

The first [37], and possibly most, studied example of this class is $TbMnO_3$, which also crystallizes in a perovskite-type structure (c.f. Fig. 10a), with Tb at the A and Mn at the B site (similar effects also occur when Tb is replaced by other rare earths, e.g. Dy or Gd). The Mn spins order around $40\,\mathrm{K}$, with an additional magnetic phase transition at $26\,\mathrm{K}$. Additionally, Tb $4f$ moment order appears below $10\,\mathrm{K}$. The magnetic transitions are well visible in magnetization and specific heat data (Fig. 13a). An electric polarization in c direction appears at the $26\,\mathrm{K}$ magnetic transition (Fig. 13b), already an indication of a "magnetically driven" FE. The value of the polarization, $800\,\mu\mathrm{C/m^2} = 0.08\,\mu\mathrm{C/cm^2}$ is very small compared to traditional FE, with typical values $> 10\,\mu\mathrm{C/cm^2}$, or also some of the MF discussed above.

By the application of a magnetic field $\|b$ the direction of the polarization can be changed from c to a (Fig. 13c/d), evidence of an enormous electromagnetic coupling [37]. The polarization flop correlates with a meta-magnetic transition, and a detailed understanding of the magnetic structure in the different phases seems to be required to address this issue, as well as the general origin of the FE in $TbMnO_3$ and related compounds. Neutron and resonant x-ray diffraction

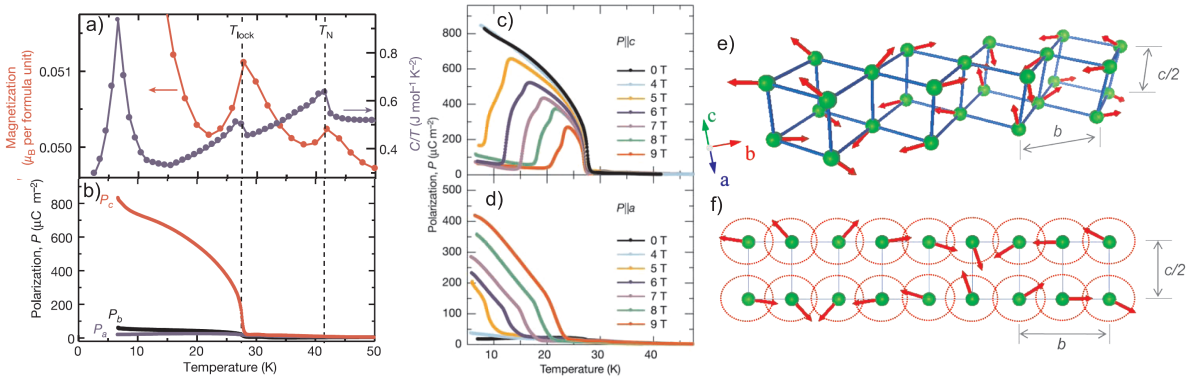


Fig. 13: *$TbMnO_3$ a/b) T dependence of magnetization, specific heat, and electric polarization. c/d) c and a component of polarization in different magnetic fields applied $\|b$. Panels a-d after [37]. e/f) Spin structure below $26\,\mathrm{K}$, with projection to bc plane [38].*

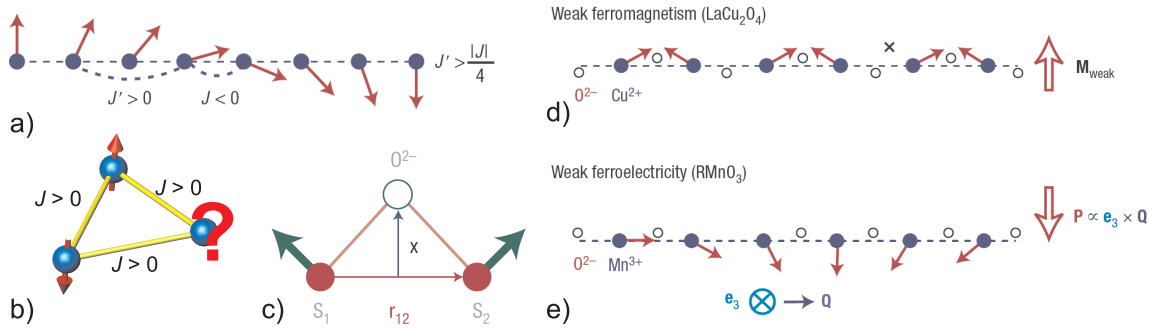


Fig. 14: *a) Spiral (cycloidal) spin order arises as classical ground state of Heisenberg spin chain with nearest neighbor FM and next-nearest neighbor AFM coupling. b) Frustration of Ising spins with AFM nearest-neighbor coupling on a triangle. c) Illustration of Dzyaloshinskii-Moriya interaction interaction (DMI) between two magnetic ions (see text). d) DMI can lead to weak ferromagnetism in antiferromagnets such as Fe_2O_3 or $LaCu_2O_4$. e) DMI can also lead to weak ferroelectricity in spiral magnets. Panels a and c-e after [39].*

data indicate that the spins initially order in an incommensurate, but almost collinear (slightly canted) amplitude-modulated structure [41, 42, 43]. The 26 K transition corresponds to the transformation into a spiral (cycloidal) spin structure, shown in Fig. 13e/f). The spins rotate around the a axis, whereas the propagation is along b. This specific spin configuration seems to be connected with a polarization along c.

The absence of an electric polarization associated with the high T spin structure and both the presence and the direction of the polarization in the low T cycloidal spin structure can be understood from symmetry-arguments on a phenomenological level [44]. While any spin ordering breaks time-reversal symmetry, an electric polarization also needs broken inversion symmetry (Fig. 7a). The spin structure above 26 K is inversion-symmetric, and therefore cannot give rise to any polarization. The cycloidal low T structure, on the other hand, is not inversion-symmetric (spatial inversion inverts the sense of direction of the rotation of the spins), allowing for an electric polarization. This is formalized using a Ginzburg-Landau free energy density approach, which in the simplest case of cubic symmetry gives a polarization [44]

$$\vec{P}(\vec{r}) = \gamma\chi_e \left[\left(\vec{M}(\vec{r}) \cdot \nabla \right) \vec{M}(\vec{r}) - \vec{M}(\vec{r}) \left(\nabla \cdot \vec{M}(\vec{r}) \right) \right], \tag{1}$$

where χ_e is the electric susceptibility. This symmetry argument shows that electric polarization can arise from spatially inhomogeneous non-centrosymmetric magnetism.

The mechanism by which such magnetic structures arise is "frustration", the existence of competing interactions. Consider for example a Heisenberg Spin chain (Fig. 14a): with only nearest-neighbor interaction $J\vec{S}_n \cdot \vec{S}_{n+1}$ the ground state is either a parallel ($J < 0$) or alternating ($J > 0$) spin-arrangement, but if there is a FM nearest-neighbor coupling $J < 0$ and a sufficiently strong AFM next-nearest-neighbor coupling $J' > -J/4$ the ground-state is a cycloid, as realized in $TbMnO_3$. Frustration and correspondingly complex spin configurations can also arise with only nearest-neighbor interactions, for example on triangular lattices with AFM coupling (Fig. 14b). Returning to the cycloid, it can be described by

$$\vec{S}(\vec{r}_n) = S_1 \vec{e}_1 cos(\vec{q} \cdot \vec{r}_n) + S_2 \vec{e}_2 cos(\vec{q} \cdot \vec{r}_n), \tag{2}$$

where the unit vectors $\vec{e}_i$, $i = 1, 2, 3$ form an orthonormal basis, $\vec{e}_3$ being the direction around which the spins rotate, and $\vec{q}$ is the propagation vector. The cycloidal ground state has $S_1 = S_2$,

$\overrightarrow{q} \perp \overrightarrow{e}_3$, and $|\overrightarrow{q}| = 2\arccos(-J'/4J)$. The spins S_i provide the magnetization M_i, and the average polarization is then obtained with Eq. (1) as

$$\langle \overrightarrow{P} \rangle = V^{-1} \int \overrightarrow{P}(\overrightarrow{r}) d^3r = \gamma \chi_e M_1 M_2 [\overrightarrow{e}_3 \times \overrightarrow{q}]. \tag{3}$$

Eq. (2) can describe also states other than a cycloidal. For example, for $\overrightarrow{q} \| \overrightarrow{e}_3$ it describes a proper screw instead: it is clear from Eq. (3) that there is no uniform polarization in this case. Also, the pure sinusoidal amplitude modulation that due to magnetic anisotropies has to replace the cycloidal ground state at sufficiently high T, as is the case in $TbMnO_3$, is described by Eq. (2), by simply setting either S_1 or S_2 to 0. Again, there is no average polarization according to Eq. (3) − consistent with the experiment. Finally, the "polarization" flop upon application of a magnetic field (Fig. 13c/d) is associated with a flop in the cycloidal spin-rotation axis, further confirming the model [45].

Intense research activity in the past few years has uncovered countless examples of materials with FE associated with spin spirals [28]. These materials have very diverse crystal structures, but Eq. (3), despite being a simplification neglecting e.g. anisotropy effects, holds in most cases. However, this phenomenological description does not address the microscopic origin of spin-spiral-based FE and in particular can not provide any estimate of the magnitude of the polarization that may be expected. A plausible microscopic mechanism [46] that at least contributes to the spin-spiral FE is the ion displacement due to the antisymmetric Dzyaloshinskii-Moriya (DM) interaction [47], which is a relativistic correction to the usual superexchange $J \overrightarrow{S}_i \cdot \overrightarrow{S}_j$. It is described by an energy term of the form

$$E_{ij}^{DM} = \overrightarrow{D}_{ij} \cdot \left(\overrightarrow{S}_i \times \overrightarrow{S}_j \right), \tag{4}$$

where $\overrightarrow{D}_{ij}$ is a material-specific vector-coefficient. For superexchange between two neighboring magnetic ions via oxygen as depicted in Fig. 14c), $\overrightarrow{D}_{ij} \| \overrightarrow{x} \propto \overrightarrow{r}_{12}$, further D_{ij} is also proportional to the spin-orbit coupling constant. As such, $|\overrightarrow{D}_{ij}|$ is a measure of local inversion-symmetry breaking, and the interaction absent when there is a center of inversion between the two magnetic ions. Discovered around 1960 [47], the DM interaction was first used to explain the weak FM observed in some primarily AFM crystals of low symmetry, such as common rust Fe_2O_3, or $LaCu_2O_4$. The situation in the latter material is sketched in Fig. 14d). Normal symmetric superexchange leads to AFM order with alternating horizontal Cu spins. However, because the oxygen ions mediating the exchange are offset from the line connecting the Cu ions, the DM term (4) is also present. In this case, $\overrightarrow{D}_{ij}$ has alternating sign, and minimization of the energy is obtained by a small canting of the spins with correspondingly alternating $\overrightarrow{S}_i \times \overrightarrow{S}_j$, yielding a net magnetic moment perpendicular to the spin chain [39].

For the spin-spiral, the reverse occurs (Fig. 14e). Here, the spin-canting is the starting point, and $\overrightarrow{S}_i \times \overrightarrow{S}_j$ has the same sign for all pairs of neighboring magnetic sites. Correspondingly, the energy can be lowered by creation of $\overrightarrow{D}_{ij}$ by displacing the oxygen ions all in the same direction perpendicular to the chain ($\overrightarrow{q}$) and the spin-rotation axis. Therefore, this "inverse DM" model predicts the same direction of the polarization as the phenomenological Eq. (3). It should be remarked that in addition to these ion displacement considerations there is also a purely electronic contribution [48]. However, comprehensive density functional calculations suggest that the ionic contribution to the FE polarization is dominant, at least in $TbMnO_3$ [49, 50]. As a microscopic model, the strength of the coupling D_{ij} and the resulting polarization can

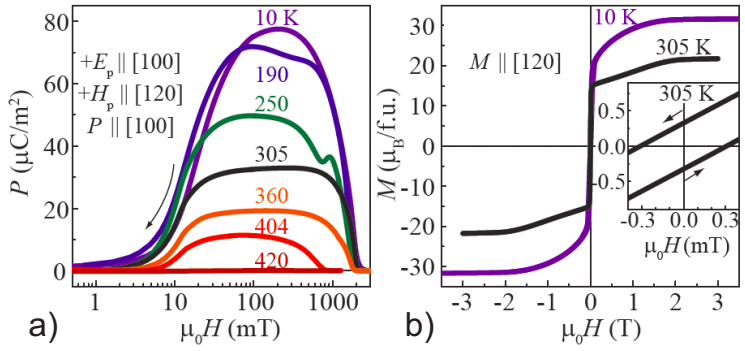


Fig. 15: *Room-temperature multiferroicity in the Z-type hexaferrite $(Ba_{0.17}Sr_{0.83})_3Co_2Fe_{24}O_{41}$ induced by magnetic fields H. a) Electric polarization P vs H at different T. b) Magnetization vs H. After [51].*

be calculated. Unfortunately, the coupling strength is set by the spin-orbit coupling, which as a relativistic effect is intrinsically very small. Correspondingly, the small polarization measured on $TbMnO_3$ is not a coincidence, but rather polarizations reaching more than $1\,\mu C/cm^2$ cannot be expected for spin-spiral based FE, values that may be too low for many applications.
Another problem with respect to potential applications until recently was the very low temperature scale, e.g. 26 K by $TbMnO_3$ (Fig. 13a), rather typical for spin-spiral FE. The basic reason for this is that the frustration inherently needed for the creation of the spin spiral also suppresses the ordering temperature. However, in recent years, spin-spiral FE with temperature scales reaching room temperature were discovered, most prominently the large family of hexaferrites with several complex structure types. Here, proper-screw type magnetic structures transform into transverse-conical spin structures upon application of a small magnetic field (see, e.g., [52]). In the hexaferrite $Ba_{0.17}Sr_{0.83})_3Co_2Fe_{24}O_{41}$ MF at room temperature has been unambiguously demonstrated (see Fig. 15 [53, 51]) and further exploration of the family holds promise of discovering compounds suitable for IT applications.

3.4 Collinear "magnetic ferroelectrics" and ferroelectricity from charge order

As the electric polarization of the spin-spiral FE is intrinsically limited by the weak spin-orbit coupling, the question arises whether there are any other mechanisms through which FE can be generated by magnetism. This is indeed the case, and I will describe here one mechanism and a corresponding example compound. Consider a chain of alternating ions (Fig. 16b). These may be different elements, or alternatively ions of one element in different valence states. The ions should be all magnetic with magnetic anisotropy providing an Ising-chain. If we have competing interactions as in Fig. 14a), but Ising instead of Heisenberg spins (i.e. we have uniaxial magnetic anisotropy), the ground state (for $J' > -J/2$) is an $\uparrow\uparrow\downarrow\downarrow$ configuration as indicated in Fig. 16b). As the magnetic exchange between neighboring ions depends on the distance between them, there is in the reverse also an exchange striction, which will tend to decrease the distance between neighboring ions with parallel spin-alignment. The result is pairs of ions corresponding to electric dipoles (c.f. Fig. 16b). It is remarkable that both the alternating-ion chain without the spins and the spin-chain ignoring the difference between the two ions are centrosymmetric. Inversion-symmetry is only broken, and FE allowed, by their combination. Because this mechanism does not depend on relativistic effects, much larger electric polarizations than in the spin-spiral FE might be expected [54].
A rather clear example of the above mechanism is provided by Ca_3CoMnO_6 [55], whose crystal structure provides for chains of alternating Co and Mn, and where Ising-magnetism with $\uparrow\uparrow\downarrow\downarrow$

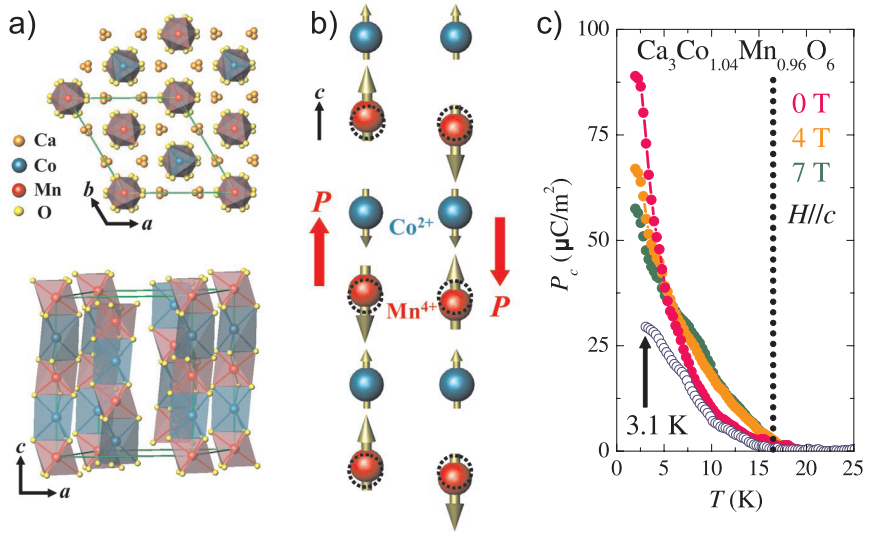


Fig. 16: *Ca_3CoMnO_6 [55]. a) Crystal structure b) Co/Mn Ising chains with two states of ↑↑↓↓ spin ordering. Ions are displaced from ideal positions (broken circles) by magnetostriction, leading to an electric polarization. c) Polarization from pyroelectric current measurements.*

spin arrangement along these chains occurs (Fig. 16a/b). As shown in Fig. 16c), pyroelectric current measurements indeed indicate a spontaneous electric polarization below the onset of magnetic order [55]. Unfortunately and contrary to the expectation, the observed polarization is even lower than in $TbMnO_3$. Likely, this exchange-striction mechanism is also the principal driving force for FE in the double-layer manganites, where however different Mn valence states provide for one of the necessary preconditions [54]. In those compounds also only small polarizations are observed.

Apart from the spin degree of freedom there are two other electronic degrees of freedom, charge and orbital. Ordering of one of these degrees of freedom may result in FE as well [54]. In particular, one can expect to obtain an electric polarization from any charge ordering (CO), the localization of charge carriers on parts of the ions in a spatially periodic fashion, that breaks spatial inversion symmetry. Although MF then is not quite as inherent as in the case of spin-based FE, the fact that the electrons that are involved in charge or orbital order inevitably also have uncompensated spins provides for a high probability of multiferroicity to result. Furthermore, because the same electrons are involved in the different orderings, a significant coupling can be expected as well. The main advantage of FE from CO, compared to "magnetic ferroelectricity", is that much higher electric polarizations can be expected.

The difficulty in obtaining FE from CO is that the Coulomb-interaction driving the CO tends to avoid non-centrosymmetric configurations. An example of how a non-centrosymmetric and thus ferroelectric CO could be obtained is given by considering charge order on a chain [54]. If there is one electron per two sites, the expectation is to obtain a chain with ions of alternating valence states. However, apart from this situation, termed "site-centered" charge order, it is also possible that the electrons localize on the bonds between the ions in a "bond-centered" charge order, a situation realized by a Peierls-distortion well-known in low-dimensional (mainly organic) compounds. Here, the ions remain equivalent, but not the bonds between them, and, as a result, there is a tendency to dimerization. Neither "site-centered" nor "bond-centered" charge order breaks inversion-symmetry by itself, but their combination does [56, 54]. The situation and resulting electric dipoles is the same as in Fig. 16b), except that the ion shifts are due to "bond-centered" charge order rather than magnetostriction (and the coloring is due to "site-centered" charge order). Such a mechanism was proposed to be active [56] in Ca-doped $PrMnO_3$ for certain intermediate Ca doping, but direct experimental confirmation of this scenario has not been obtained, not least due to a too large conductivity of (Pr,Ca)MnO_3.

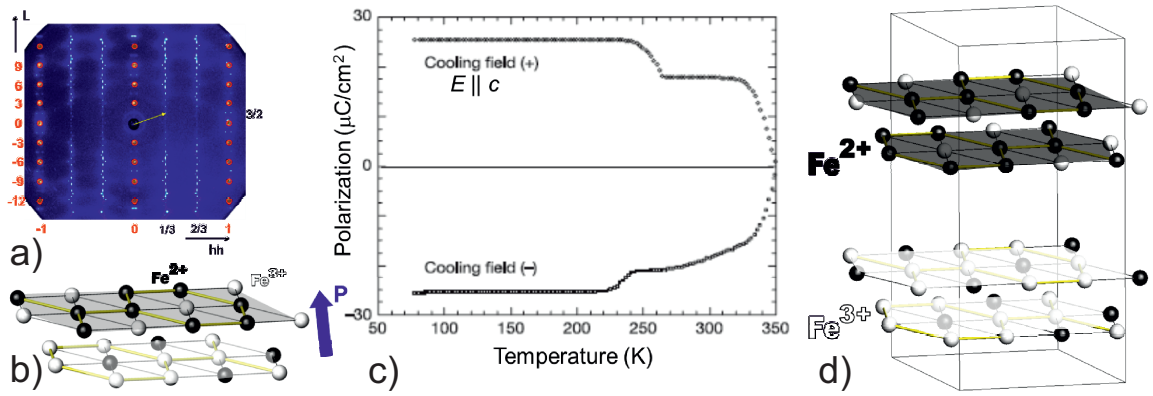


Fig. 17: *Ferroelectric charge order CO in $LuFe_2O_4$? a) CO reflections by high-energy x-ray diffraction. b) Model, proposed in 2005 [57], of the $Fe^{2+/3+}$ CO in Fe/O bilayers, the basic structural subunit. Different average valence of the two layers makes the bilayer polar, indicated by the arrow. c) Pyroelectric current measurements [57] indicated a large ferroelectric polarization. d) Recently revised model of the CO, based on the first structure refinements [58]: The bilayers are charged rather than polar, excluding ferroelectricity due to CO.*

Unambiguous confirmation of FE due to CO is elusive also in other compounds so far, and there are as yet only few examples, none of them absolutely clear. We first discuss an example that scattering methods recently showed is rather a non-example: The material most often cited as the prototypical example of CO-based FE (see, e.g., [28]) is $LuFe_2O_4$, which has a layered crystal structure with Fe/O bilayers as the the electronically active subunit. The Fe ions have an average valence of $2.5+$, so a Fe^{2+}/Fe^{3+} charge ordering can be expected. Superstructure reflections that can be attributed to charge order do indeed appear slightly above room temperature (Fig. 17a). The location of these reflections at "$(h, h) = (\frac{1}{3}, \frac{1}{3})$" and equivalent is consistent with the charge configuration within a bilayer indicated in Fig. 17b), first proposed by Ikeda and coworkers in 2005 [57]. Because one of the two layers is rich in Fe^{2+} and the other in Fe^{3+} (as indicated by a shading of the layers), the bilayer would become polar by such a CO. In accordance with this polar CO model, pyroelectric current measurements (Fig. 17c) suggested a spontaneous polarization, also up to slightly above room temperature [57]. Although these polar bilayers were generally accepted in the extensive literature on $LuFe_2O_4$, a direct proof for or against was lacking. Only very recently, the charge-ordered crystal structure could be refined based on single-crystal x-ray diffraction data measured on an almost-mono-domain crystal. The Fe^{2+}/Fe^{3+} arrangement deduced from the refinement, which is supported also by x-ray magnetic circular dichroism and neutron diffraction, contains bilayers that are charged rather than polar (Fig. 17d), excluding CO-based FE in $LuFe_2O_4$ [58].

A remarkable, *if verified*, example of CO-driven FE would be classical magnetite Fe_3O_4, the oldest magnetic material known to mankind and also the classical (1939) example of a CO transition. Despite decades of research, the very complex CO structure of magnetite has eluded a determination until recently, when Senn and coworkers achieved a refinement based on high-energy x-ray diffraction on a tiny ($40\,\mu$m) almost-mono-domain crystal [59]. The refined structure, consistent with density-functional calculations [60] is indeed polar, which leaves only to show that it can actually be switched by an electric field. The latter is complicated by residual conductivity and associated "relaxor-like characteristics", indicated by dielectric spectroscopy [61].

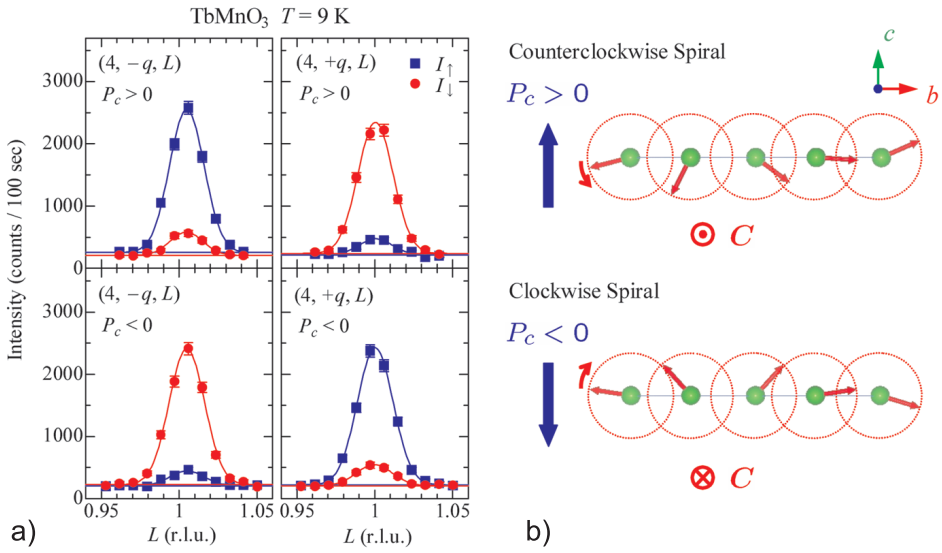


Fig. 18: *Electric-Field control of magnetic state in $TbMnO_3$ [38]. a) Polarized neutron scattering dependence on Polarization, selected by voltage bias on cooling. b) correspondence of spin-spiral helicity and polarization.*

3.5 Electric-field control of magnetism and multiferroic heterostructures

Indications for magnetoelectric coupling have been observed for many MF within the different "classes" described above. For the spin-spiral FE in particular, control of the electric polarization by a magnetic field has been shown in a dramatic way already in the first paper [37] (Fig. 13c/d). Given the mechanism by which the polarization arises directly from the spin structure, this is hardly surprising. However, for many applications such as the MF-RAM (Fig. 9) it is more important to control the magnetization by an electric field. Demonstrations of magnetoelectric coupling in this direction are much more rare. Still, the DM interaction as an energy contribution Eqn. (4) works both ways, as we have seen also from the example of weak FM.

In the spin-spiral FE there is generally no spontaneous magnetization that can be switched, but on the other hand reversing the sense of rotation along the spiral reverses the associated FE polarization (Fig. 18b), therefore switching the polarization by an applied electric field should be able to revert the sense of rotation of the spins. The technique most suited to measure the sense of rotation is polarized neutron scattering, because the intensity of magnetic reflections will be different for the neutron spins parallel or antiparallel to the chirality of the spin-spiral. Measurements on $TbMnO_3$ (Fig. 18a) indeed showed different sense of rotations of the spins after cooling to the FE phase in a positive or negative electric field [38].

Whereas strong magnetoelectric coupling is inherent in the spin-spiral FE, this is not the case for MF in general, as we have seen on the example of $BiMnO_3$. Surprisingly, electric-field control of magnetism is possible in $BiFeO_3$ [62], despite of this material being in the same "class" of MF as $BiMnO_3$. To understand why, we need to consider the magnetic structure [36]: it is based on a G-type AFM, where each Fe site spin has six antiparallel nearest neighbors, but overlaid on this are a spin-canting and long-period cycloidal modulation (Fig. 19b). The cycloidal modulation in this case does not arise from magnetic frustration, but rather is a consequence of the DM interaction (4). Indeed, it was long known that apart from creating weak FM, the DM interaction can also lead to cycloidal spin configurations. The mechanism is exactly as discussed above and sketched in Fig. 14e), except that the off-centering of the oxygen ions is the origin rather than the effect. In $BiFeO_3$, this off-centering is due to the electric polarization, which in turn is driven by the "lone-pair" mechanism. This compound is therefore the opposite to $TbMnO_3$: rather than FE due to spin-spiral, we have spin-spiral due to FE. Different direc-

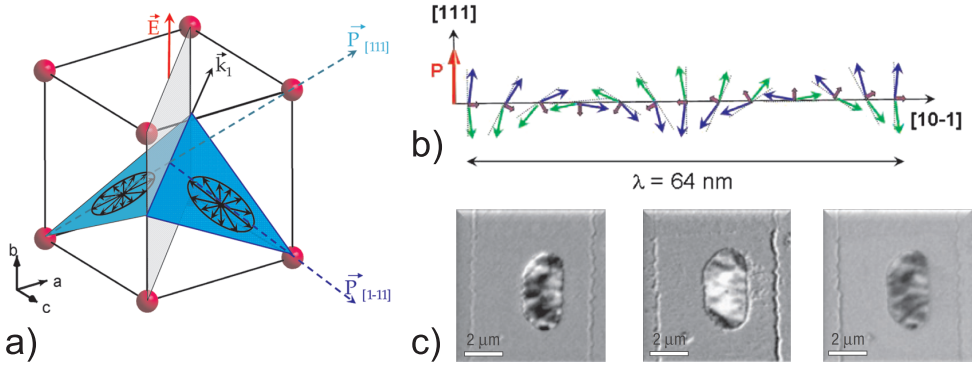


Fig. 19: *Electric-Field control of magnetic state in $BiFeO_3$. a) sketch of spiral domains and relation-ship with polarization [62]. Domains can be selected by external electric field. b) in one domain, spin structure is given by superposition of G-type antiferromagnetic order, spin canting, and long-period spiral [62]. c) XMCD-PEEM images [63] showing magnetic contrast of a CoFe-layer, whose magnetization is switched using exchange-bias to an adjacent $BiFeO_3$ layer, the magnetism of which is controlled by an electric field. Images left to right are obtained by two consecutive electrical field switches.*

tions of the polarization will correspond to different possible directions (domains) of the spin spiral (Fig. 19a), providing a mechanism for switching of magnetic domains by an electric field, as demonstrated by Lebeugle and coworkers [62].

Because all this works even at room temperature, $BiFeO_3$ appears to be a most promising candidate for building MF-RAM cells (Fig. 9) and similar applications. What is needed next is a mechanism, such as exchange bias (see lecture E5), through which the magnetic domain configuration of $BiFeO_3$ is linked to the ferromagnetic domain structure of an adjacent soft FM. This was successfully demonstrated by Chu and co-workers [63], who built a heterostructure with $BiFeO_3$ and CoFe as the soft FM (see also [64]. In Fig. 19c) three XMCD-PEEM images of this device, with a small CoFe island in the center, are shown. The contrast is given by different magnetic polarities. Between neighboring images, an electric field is applied, first in one, then in the other direction. The clearly visible switching of the CoFe magnetization by the electric field was achieved at room temperature.

In multiferroic devices such as the MF-RAM (Fig. 9) the materials occur not in the bulk form, but in (patterned) heterostructures. In this case interface effects can be of crucial importance. A good example is the use of the exchange-bias effect between a FM and an AFM FE, e.g. for a CoFe-$BiFeO_3$ heterostructure as discussed just above (c.f. Fig. 19c). But with interface effects it is even possible to create composite multiferroics out of materials that are all non-multiferroic intrinsically. One approach (for reviews see, e.g., [65, 66]) uses strain at the interface to couple a FE material (via the piezoelectric effect) to a FM material (via magnetostriction). This works best with FE with large piezoelectric coefficents (in particular $Pb(Zr,Ti)O_3$ or PZT) and FM with large magnetostriction. In such systems, magnetoelectric coupling coefficients more than two orders of magnitude larger than in single-phase MF have been observed. However, while the concept of strain-coupling is simple, the detailed modeling is difficult, and so is experimental reproducibility, given strong influence of defects, domains etc.

A more recent approach uses the strong sensitivity of correlated oxides to charge carrier density, employing FE gate oxides. The two polarization states of the gate, tunable by low voltages, then correspond to two different effective doping levels in the magnetically active correlated oxide.

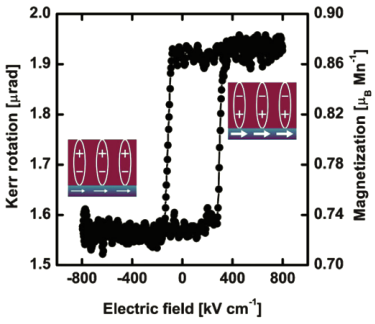


Fig. 20: *Magnetoelectric hysteresis-loop showing the magnetic response (Kerr-rotation $\propto$ M) as a function of applied electric field at* $100\,\mathrm{K}$ *of a PZT(*$250\,\mathrm{nm}$*)-*$La_{0.8}Sr_{0.2}MnO_3$*(*$4\,\mathrm{nm}$*) heterostructure grown on a* $SrTiO_3$*(001) substrate. Insets represent the magnetic and electric states of the two layers. After [67].*

For the latter, doped perovskite manganites, such as $La_{1-x}Sr_xMnO_3$, which are characterized by competing phases and an intricate interplay of magnetism, structure, and electron transport [68], can be employed. For an appropriate composition, the additional effective doping by the gate can trigger a transition to a phase with different magnetic properties. Following this recipe, Molegraaf and coworkers [67] have recently demonstrated a large magnetoelectric coupling in a PZT-$La_{0.8}Sr_{0.2}MnO_3$ bilayer heterostructure (Fig. 20). The effect shown is much larger than in single-phase multiferroics and approaches values of strain-coupled multiferroic heterostructures, while being of purely electronic origin [69]. Similarly, an electronic reconstruction at the interface between a FM and a FE can induce a purely magnetoelectric coupling, as recently demonstrated for $BaTiO_3$ and Fe or Co [70].
For the realization of an MF-RAM (Fig. 9) and further spintronic applications it is also necessary to consider the transport properties of multiferroic heterostructures, i.e. multiferroic tunnel-junctions. Such a tunnel junction depending both on electric and magnetic states for a total of four memory states has already been demonstrated [71]. There is currently much research being conducted in this field closest to application, and promising results can be expected.

4 Concluding remarks

On the way towards non-volatile memory applications actually on the market, phase change materials are further advanced than multiferroics. In fact, a 512 Mbit PC-RAM is already used as a replacement of slower flash in some mobile phones [72]. However, the relatively low storage capacity compared to the currently available 64 Gbit flash-based USB sticks suggests scalability problems and additionally the high operating current densities force device architectures that put the PC-RAM to a severe cost-disadvantage compared to Flash. These problems will have to be overcome if PC-RAM is to become more than a niche product. Aiming towards an universal memory (Sec. 1), new routes need to be pursued in research. One such route, going from single phases to heterostructures, is starting to be pursued, with promising results [73]: in a multilayer $(GeTe)_2$-$(Sb_2Te_3)_4$ system, switching currents considerably lower and lifetimes orders of magnitude higher than for $Ge_2Sb_2Te_5$ have been observed.
For multiferroics, on the other hand, the room-temperature switchability of a magnetic layer, crucial for the MF-RAM, has been shown using $BiFeO_3$ [63, 64], and also on heterostructures such as Pb(Zr,Ti)O_3-$La_{0.8}Sr_{0.2}MnO_3$ [67], though the issue of the (presumably fast) time scales of the switching have to be addressed experimentally. However, so far a fully functioning MF memory cell has yet to be demonstrated, and from there to an integrated circuit it is still a way to go. All this suggests that it will still be some time before we can expect to buy devices that use MF-RAM. A large contrast to the PCM with their rather common mechanism

and clear location on a "treasure map" (Fig. 5a) is that there are many potential ways towards multiferroicity. Currently rather classical $BiFeO_3$ with its lone-pair ferroelectricity seems to have the lead, but good progress is also being made recently with spin-spiral based ferroelectrics such as hexaferrites, and artificial heterostructures show promise as well. This richness of materials, effects, and mechanisms in multiferroics provides for some unpredictability, but also for chances in the medium and long term. Of course, in the long term, some of the various more exotic concepts mentioned in Sec. 1, that could not be covered in detail, might prevail.
In any case, research on a materials or heterostructure level, will continue for both PCM and multiferroics to contribute critically to the eventual realization of integrated devices. I hope that this lecture also showed how the application of a large variety of scattering methods is crucial to make progress in understanding these PCM and multiferroics. Although the scattering examples shown are all from bulk systems, I would like to stress that scattering methods are also indispensable for the detailed understanding of heterostructures, the last step short of a device.

Acknowledgment

I gratefully acknowledge R. P. Hermann for contributing graphics used in Sec. 2, and A. Weber for proof-reading of the lecture notes.

References

[1] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[2] L. DiCarlo *et al.*, Nature **460**, 240 (2009).

[3] F. Schwierz, Nat. Nanotech. **5**, 487 (2010).

[4] R. Waser, R. Dittmann, G. Staikov, and K. Szot, Adv. Mater. **21**, 2632 (2009).

[5] L. Bogani and W. Wernsdorfer, Nat. Mater. **7**, 179 (2008).

[6] P. Fromherz, A. Offenhausser, T. Vetter, and J. Weis, Science **252**, 1290 (1991).

[7] M. Wuttig and N. Yamada, Nat. Mater. **6**, 824 (2007).

[8] I. Friedrich, V. Weidenhof, S. Lenk, and M. Wuttig, Thin Solid Films **389**, 239 (2001).

[9] W. Welnic and M. Wuttig, Materials Today **11**, 20 (2008).

[10] N. Yamada *et al.*, J. Appl. Phys. **69**, 2849 (1991).

[11] G. Atwood, Science **321**, 210 (2008).

[12] S. R. Ovshinsky, Phys. Rev. Lett. **21**, 1450 (1968).

[13] C. Kim *et al.*, Appl. Phys. Lett. **94**, 193504 (2009).

[14] K. Shportko *et al.*, Nat. Mater. **7**, 653 (2008).

[15] L. Pauling, *Nature of Chemical Bond* (Cornell Univ. Press, New York, 1939).

[16] D. Lencer *et al.*, Nat. Mater. **7**, 972 (2008).

[17] S. Kohara *et al.*, Appl. Phys. Lett. **89**, 201910 (2006).

[18] J. Hegedus and S. R. Elliott, Nat. Mater. **7**, 399 (2008).

[19] T. Matsunaga *et al.*, Adv. Funct. Mat. **21**, 2232 (2011).

[20] H. Schmid, Ferroelectrics **162**, 317 (1994).

[21] N. A. Spaldin, M. Fiebig, and M. Mostovoy, J. Phys.: Condens. Matter **20**, 434203 (2008).

[22] M. Bibes and A. Barthélémy, Nat. Mater. **7**, 425 (2008).

[23] J. Nogués and I. K. Schuller, J. Magn.Magn. Mater. **192**, 203 (1999).

[24] N. A. Hill, J. Phys. Chem. B **104**, 6694 (2000).

[25] D. I. Khomskii, J. Magn. Magn. Mater. **306**, 1 (2006).

[26] S. Piccozzi and C. Ederer, J. Phys.: Condens. Matter **21**, 303201 (2009).

[27] D. I. Khomskii, Physics **2**, 20 (2009).

[28] K. F. Wang, J.-M. Liu, and Z. F. Ren, Adv. Physics **58**, 321 (2009).

[29] R. Seshadri and N. A. Hill, Chem. Mater. **13**, 2892 (2001).

[30] T. Kimura *et al.*, Phys. Rev. B **67**, 180401(R) (2003).

[31] J. Y. Son and Y.-H. Shin, Appl. Phys. Lett. **93**, 062902 (2008).

[32] J. Wang *et al.*, Science **299**, 1719 (2003).

[33] D. Lebeugle, D. Colson, A. Forget, and M. Viret, Appl. Phys. Lett. **91**, 022907 (2007).

[34] Note that ferroelectricity in ideal stoichiometric $BiMnO_3$ has recently become a very controversial issue, with antiferroelectricity as an alternative. For a summary of different evidences see I. V. Solovyev and Z. V. Pchelkina, New J. Phys. **10**, 073021 (2008).

[35] G. Catalan and J. F. Scott, Adv. Mater. **21**, 2463 (2009).

[36] I. Sosnowska, T. Peterlin-Neumaier, and E. Steichle, J. Phys. C:Solid State Phys. **15**, 4835 (1982).

[37] T. Kimura *et al.*, Nature **426**, 55 (2003).

[38] Y. Yamasaki *et al.*, Phys. Rev. Lett. **98**, 147204 (2007).

[39] S.-W. Cheong and M. Mostovoy, Nat. Mater. **6**, 13 (2007).

[40] Y. Tokura and S. Seki, Adv. Mater. **21**, 1 (2009).

[41] M. Kenzelmann *et al.*, Phys. Rev. Lett. **95**, 087206 (2005).

[42] J. Voigt *et al.*, Phys. Rev. B **76**, 104431 (2007).

[43] S. B. Wilkins *et al.*, Phys. Rev. Lett. **103**, 207602 (2009).

[44] M. Mostovoy, Phys. Rev. Lett. **96**, 067601 (2006).

[45] N. Aliouane *et al.*, Phys. Rev. Lett. **102**, 207205 (2009).

[46] I. A. Sergienko and E. Dagotto, Phys. Rev. B **73**, 094434 (2006).

[47] I. Dzyaloshinsky, J. Phys. Chem. Solids bf 4, 241 (1958); T. Moriya, Phys. Rev. **120**, 91 (1960).

[48] H. Katsura, N. Nagaosa, and A. Balatsky, Phys. Rev. Lett. **95**, 057205 (2005).

[49] H. J. Xiang, S.-H. Wei, M.-H. Whangbo, and J. L. D. Silva, Phys. Rev. Lett. **101**, 037209 (2008).

[50] A. Malashevich and D. Vanderbilt, Phys. Rev. Lett. **101**, 037210 (2008).

[51] S. H. Chun *et al.*, arXiv:1111.4525.

[52] M. Soda *et al.*, Phys. Rev. L **106**, 087201 (2011).

[53] Y. Kitagawa *et al.*, Nat. Mat **9**, 797 (2010).

[54] J. van den Brink and D. I. Khomskii, J. Phys.: Condens. Matter **20**, 434217 (2008).

[55] Y. J. Choi *et al.*, Phys. Rev. Lett. **100**, 047601 (2008).

[56] D. V. Efremov, J. Van den Brink, and D. I. Khomskii, Nat. Mater. **3**, 853 (2004).

[57] N. Ikeda *et al.*, Nature **436**, 1136 (2005).

[58] J de Groot *et al.*, arXiv:1112.0978v1.

[59] M. S. Senn, J. P. Wright, and J. P. Attfield, Nature **481**, 173 (2012).

[60] K. Yamauchi, T. Tetsuya, and S. Piccozzi, Phys. Rev. B **79**, 212404 (2009).

[61] F. Schrettle *et al.*, Phys. Rev. B **83**, 195109 (2011).

[62] D. Lebeugle *et al.*, Phys. Rev. Lett. **100**, 227602 (2008).

[63] Y.-H. Chu *et al.*, Nat. Mater. **7**, 478 (2008).

[64] J. T. Heron *et al.*, Phys. Rev. Lett. **107**, 217202 (2011).

[65] R. Ramesh and N. A. Spaldin, Nat. Mater. **6**, 21 (2007).

[66] C. A. F. Vaz, J. Hoffman, C. H. Ahn, and R. Ramesh, Adv. Mat **22**, 2900 (2010).

[67] H. J. A. Molegraaf *et al.*, Adv. Mater **21**, 3470 (2009).

[68] A. P. Ramirez, J. Phys.: Condens. Matter **9**, 8171 (1997).

[69] C. A. F. Vaz *et al.*, Phys. Rev. Le **104**, 127202 (2010).

[70] S. Valencia *et al.*, Nat. Mater. **10**, 753 (2011).

[71] M. Gajek *et al.*, Nat. Mater. **6**, 296 (2007).

[72] http://www.theregister.co.uk/2010/04/30/samsung_pram_ships/.

[73] R. E. Simpson *et al.*, Nat. Nanotech. **6**, 501 (2011).

E 5 Thin Film Heterostructures

E. Kentzinger, S. Pütter
Jülich Centre for Neutron Science (JCNS)
Forschungszentrum Jülich GmbH

1 Introduction

Thin film heterostructures are composed of layers of different materials and of thicknesses in the nanometer range. The heterostructures can be laterally homogeneous or may present lateral density fluctuations on mesoscopic length scales, i. e. from nanometers to micrometers. Reduced dimensions, reduced dimensionalities and the presence of interfaces lead to very interesting phenomena. Some of them will be described in the following. Thin film heterostructures are present in many areas of science and technology e.g. in soft matter, life science, or information technology.

A lot of **soft matter** systems are investigated as thin films. For examples ordering effects at the liquid-solid, liquid-liquid and liquid-air interfaces might appear, the interfaces inducing a break of rotational invariance of the liquid. Those effects are most pronounced when the compositional units of the liquid (surfactant, lipid molecule, etc.) are, themselves, not spherical [1–3]. Another area of investigation is polymer thin films, driven by enormous technological potential and the intellectually challenging academic problems associated with them [4]. Among them are block copolymer (BCP) thin films. BCPs consist of two or more chemically different polymer chains joined covalently at their ends. Due to the repulsive interactions and small entropy of mixing, dissimilar blocks tend to microphase separate into well-ordered arrays of spherical, cylindrical, gyroid or lamellar microdomains depending on the volume fractions of the blocks and the degree of microphase separation. In contrast to the bulk, the morphology of BCP thin films can be strongly influenced by surface and interfacial energies as well as the commensurability between the film thickness and the period of the microdomain morphology. With decreasing film thickness these parameters become increasingly important in defining the morphology.

In the **life science** area, nanostructured thin films and surfaces play an important role [5]. For example, all biological cells are enclosed by membranes that define their boundaries and regulate their interaction with the environment. The biological membrane consists of assemblies of phospholipid and protein molecules. The phospholipid molecules form a continuous double layer, or bilayer, which acts as a barrier to water-soluble molecules and provides the framework for the incorporation of the protein molecules. The proteins are receptors for hormones, drugs, medicaments, antibodies and regulate the ion-permeability of the membrane. The intrinsic complexity of cell membrane systems often precludes direct access to these features, thus driving the development of simpler model systems that are more amenable to a detailed characterization. One generic approach involves supports for the stabilization of biomimetic membranes [6, 7].

Since the 1950's **magnetic** films have been used in computer hard disk drives. From the 1980's, the possibility to prepare magnetic multilayers, materials confined at the nanometer length scale along one dimension has led to a new class of systems with extraordinary physical properties, on the fundamental point of view as well for applications in information technology. Physical effects like the interlayer exchange coupling, giant magnetoresistance, tunnel magnetoresistance or the current induced magnetization switching [8, 9] are based on the fact that not only the charge but also the spin of the electron can be used for information storage and processing. The development of lithography [10, 11] and self organization methods [12, 13] has led to systems of even lower dimensionalities like arrangements of lines (1D) or dots (0D). New effects might be expected if the size of the element or the period of the arrangement is smaller than some characteristic length scale of the material, like the wavelength of the electrons at the Fermi level, the spin diffusion length of the electrons, or the width of the domain wall [10, 14]. For example,

the magnetization reversal process is strongly modified in a magnetic bubble of diameter such that the domain formation is excluded. Typically, those sizes stem from several nanometers to several tens of nanometers. On the point of view of applications, magnetic nanostructures are the ultimate building blocks of components such as random access memories or patterned recording media. The path towards the miniaturization of such components and the proximity of the superparamagnetic limit make the magnetic interaction between the components becoming a more and more important parameter that has to be understood and controlled [15].
The active search for new materials that would allow even higher sensitivity and controllability than in the above mentioned ones based on simple transition metals is under way. **Transition metal oxides** (TMO) are particularly attractive, since there is a plethora of isostructural materials with a wide variety of magnetic and electronic properties, which can be built into heterostructures [16]. In heterostructures composed of different TMO, the disruption introduced even by an ideal interface can drastically upset the delicate balance of the competing interactions among electronic spins, charges and orbitals, leading to a range of exotic phenomena, including interfacial conduction, magnetism and superconductivity [17–21].
Neutron and X-ray scattering are methods of choice for the investigation of thin film heterostructures. Specular reflectivity gives access to density profiles on the nanometer length scale. Off-specular scattering and grazing incidence small angle scattering permit, depth resolved, the investigation of lateral fluctuations respectively at the micrometer and nanometer length scales [22, 23]. Resonant X-ray scattering, in the hard and soft energy regions, permit, element sensitive, the investigation of spin, charge and orbital orderings in thin film heterostructures [24–26]. In the following, we present four examples of investigation of thin film heterostructures, coming from the scientific fields cited above.

2 Crystallization of micelles at chemically terminated interfaces

In this part we present an investigation of the ordering of spherical micelles in an aqueous solution in the immediate vicinity of a flat solid substrate. The authors of this study showed for the first time that, by a combination of neutron reflectivity and grazing incidence small angle neutron scattering (GISANS), the structure of buried adsorbed layers can be resolved in three dimensions over length scales reaching from the nm to the μm regime [27].
The studied sample is made out of triblock copolymers (so called “poloxamer”) composed of a central hydrophobic polymer chain flanked by to hydrophilic chains. In aqueous solution those poloxamers form spheres with the hydrophilic “head” regions in contact with the surrounding solvent, sequestering the hydrophobic single tail region in the micelle center (Fig. 1).

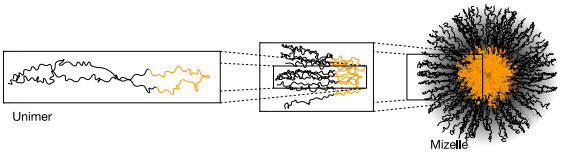


Fig. 1: *Schematics of the formation of a micelle in aqueous solution out of a triblock copolymer. After [28].*

For a critical micelle volume fraction a percolation transition occurs. It is proposed that the percolated state is metastable, and the micelles crystallize when the system is disturbed, e. g. by the application of shear force or at a solid interface. For the neutron scattering experiments presented here, the poloxamer was diluted in deuterated water in order to maximize the scattering contrast.

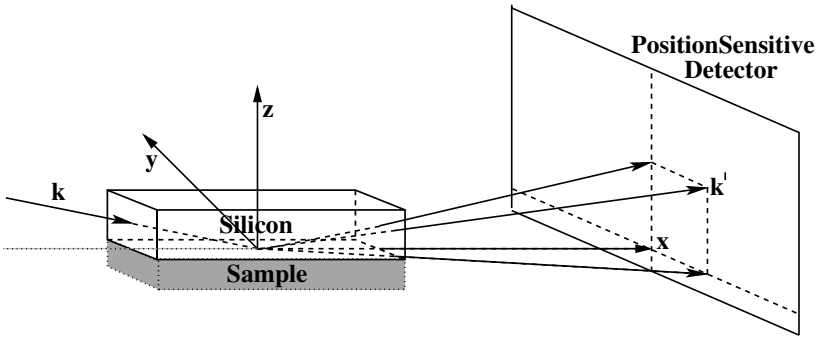


Fig. 2: *Scattering geometry for the reflectometry and GISANS studies. After scattering at the solid-liquid interface the neutrons are detected by a position sensitive detector. After [27].*

Two single crystalline polished silicon disks were prepared as solid interfaces, that were treated resulting in a hydrophilic termination for the first one and a hydrophobic termination for the second one.
The neutron scattering experiments were performed on the ADAM reflectometer at the Laue Langevin Institut (ILL). Fig. 2 illustrates the scattering geometry. A monochromatic neutron beam (marked by arrows) enters a block of silicon from the side and becomes scattered at the solid-liquid interface. The intensity is registered on a position sensitive detector.
Fig. 3 shows the true specular reflectivity intensity taken with the sample in the liquid phase at 295 K at the hydrophilic (left panel) and hydrophobic (right panel) interface. Both data sets are compared to the Fresnel reflectivity R_F. At $Q_z = 0.04$ Å^{-1}, the intensity is increased for the hydrophilic interface, showing layers of adsorbed micelles. For the hydrophobic interface the reflectivity drops faster than R_F and no peak is visible. This shows a huge interface roughness and no absorption of micelles takes place (see scheme in Fig. 4).

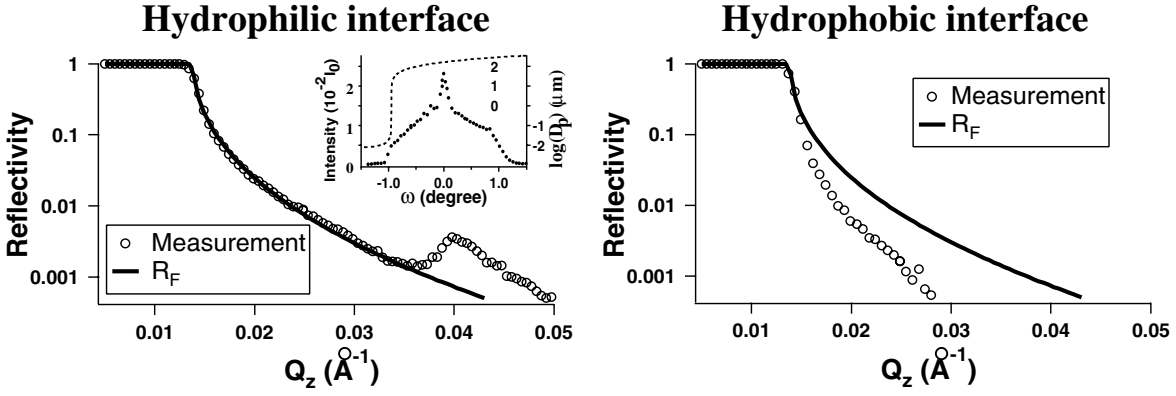


Fig. 3: *Reflectivity data taken with the sample in the liquid phase in contact with the hydrophilic (left panel) and the hydrophobic (right panel) interface compared to the Fresnel reflectivity R_F. For the hydrophilic interface a reflection from adsorbed micelle layers is visible at $Q_z = 0.04$ Å^{-1}. After [27].*

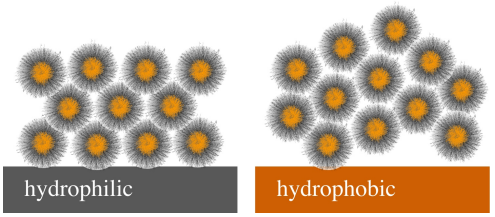


Fig. 4: *Schematic representation of the out-of-plane ordering of spherical micelles at the hydrophilic (left panel) and hydrophobic (right panel) interface. After [28].*

Fig. 5 shows GISANS data taken at an incident angle of 0.30°, slightly larger than the total reflection edge (0.25°), resulting in a penetration depth of the neutrons in the liquid of 40 μm. The total sample thickness was 1 mm. The panels on top [5(a)-(c)] and bottom [5(d)-(f)] show data for the hydrophobic and the hydrophilic interfaces, respectively. All data are corrected for refraction effects, i. e. $Q_z = \sqrt{Q_{z,m}^2 - 4\pi(\sigma_L - \sigma_{Si})}$, where $Q_{z,m}$ denotes the measured momentum transfer in the z direction and σ_L and σ_{Si} are the coherent scattering length densities of the liquid and the silicon, respectively.

For the left panels the sample is in the liquid phase at a temperature of 295 K. A ring of increased intensity with radius 4×10^{-2} Å^{-1} is visible for both interfaces, corresponding to a mean distance between micelles of 160 Å. At the hydrophilic interface sharp Bragg reflections become visible when the sample is heated to 298 K. The dots in the left panel of Fig. 6 show the reciprocal lattice for a cubic close packed (ccp) structure with the [111] direction perpendicular to the interface. When the perpendicular axis [1-10] and [11-2] are randomly oriented around the unique direction [111], the dots become smeared out along the solid and dashed lines. This structure corresponds to a two-dimensional powder. Note that in this scattering geometry $|\mathbf{k}| \approx 100 \times |Q^{111}|$, and thus the Ewald sphere cuts through these powder rings (Fig. 6). The diffraction peaks visible in Fig. 5 are shown as circles in Fig. 6, while reciprocal lattice point not probed by the experiment are shown as dashed line. The intensities of the diffraction peaks decrease

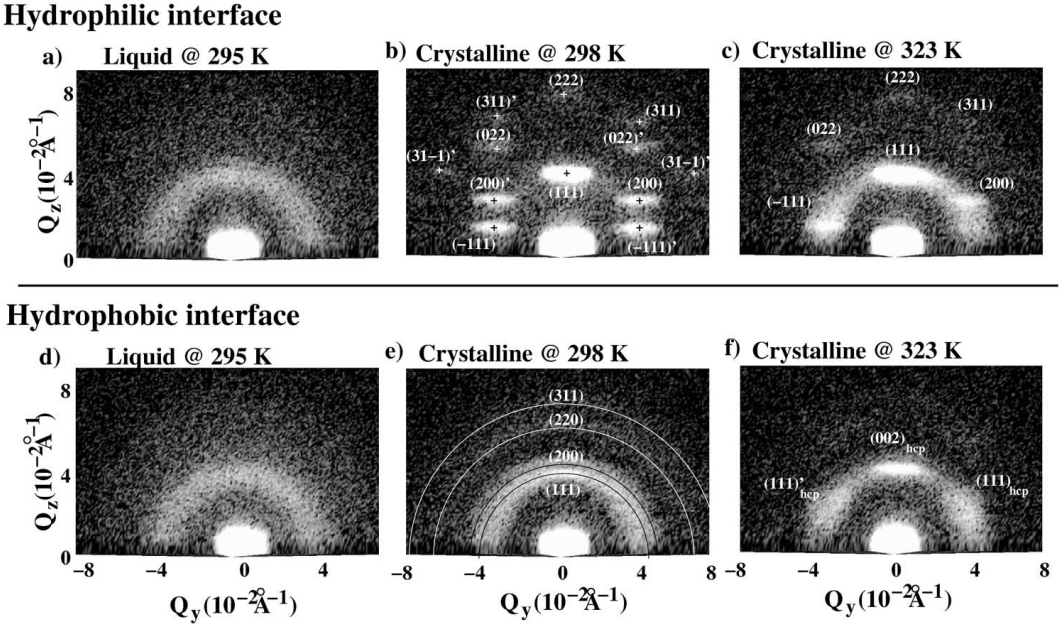


Fig. 5: *GISANS data taken at a hydrophilic (a)-(c) and a hydrophobic (d)-(f) interface in the liquid phase at 295 K (a), (d) and in the crystalline phase at 298 K (b), (e) and 323 K (c), (f). After [27].*

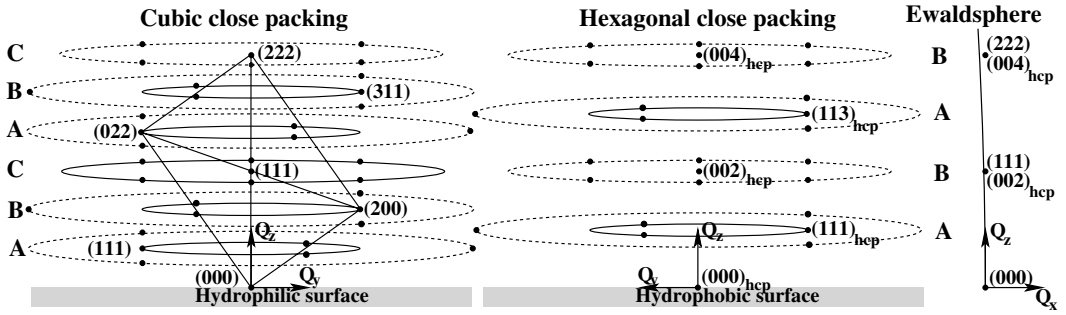


Fig. 6: *Reciprocal lattice for a cubic (ccp) and a hexagonal (hcp) close packing. The dots mark the positions of the reciprocal lattice points. For a two-dimensional powder with the unique axis along Q_z perpendicular to the interface these points are smeared out into circles. On the right side the positions of the Bragg reflections in the z direction are shown with respect to the Ewald sphere. After [27].*

by three factors: first, an increasing diameter of the powder circles; second, the drop-off of the form factor of the micelles; third, an increasing distance of the reciprocal lattice points from the Ewald sphere (Fig. 6, right side). Only the smearing of the reflections due to mosaicity or due to the size of the crystallites contributes to the detected intensity.

The data for the hydrophobic interface at 298 K feature four diffraction rings. Their absolute positions are explained by a three-dimensional polycrystalline cubic phase centered structure with a lattice constant of 295 Å. When the temperature is increased to 323 K [5(c)], the sample at the hydrophilic interface becomes textured, and a rotational orientation develops with respect to the common axis. At these temperatures different reflections emerge also for the hydrophobic interface. These are explained by a hexagonal close packing (hcp) (Fig. 6) again with random in-plane orientation. The smearing of the $(111)_{\mathrm{hcp}}$ shows the presence of staple faults as the hcp develops from a cubic lattice.

In summary, the effect of different surface potentials on the crystallization of spherical micelles at flat solid interface has been investigated by neutron reflectometry and GISANS. It is found that micelles with a hydrophilic shell crystallize in a cubic face centered (lattice constant 29.5 nm) at a hydrophilic interface. When crossing the percolation limit, first a 2D powder with a ccp structure is formed distinguished by an unique axis perpendicular to the interface. For high temperatures a rotational alignment of crystallites around the common axis develops. For the hydrophobic interface crystallization neglects the interface and a three-dimensional powder forms. It transforms into a hcp structure with increasing temperature.

In conclusion, it is shown that the structural arrangement of spherical micelles at flat surfaces is strongly affected by the surface potential. Furthermore, GISANS was found to be an excellent tool to gain information on the near surface structure of particles that are adsorbed from the solution at the solid interface.

3 Structural characterization of sparsely tethered bilayer lipid membranes

Here we show an investigation of a biomimetic membrane system, illustrated in Fig.7, that is chemically tethered to a gold support through a tether lipid. Tethered membranes are systems designed for the incorporation of membrane-associated proteins. In order to be useful as a biomembrane model, such membranes need to retain their fluid in-plane organization and at the same time remain separated from the supporting solid interface by a molecularly thin hydration layer. In order to create space for hydration, the tether lipid is co-adsorbed with a smaller "backfiller" molecule (see Fig. 7). The resulting "sparsely-tethered membrane" has been characterized by ellipsometry, electrochemical impedance spectroscopy and neutron reflectometry [7].

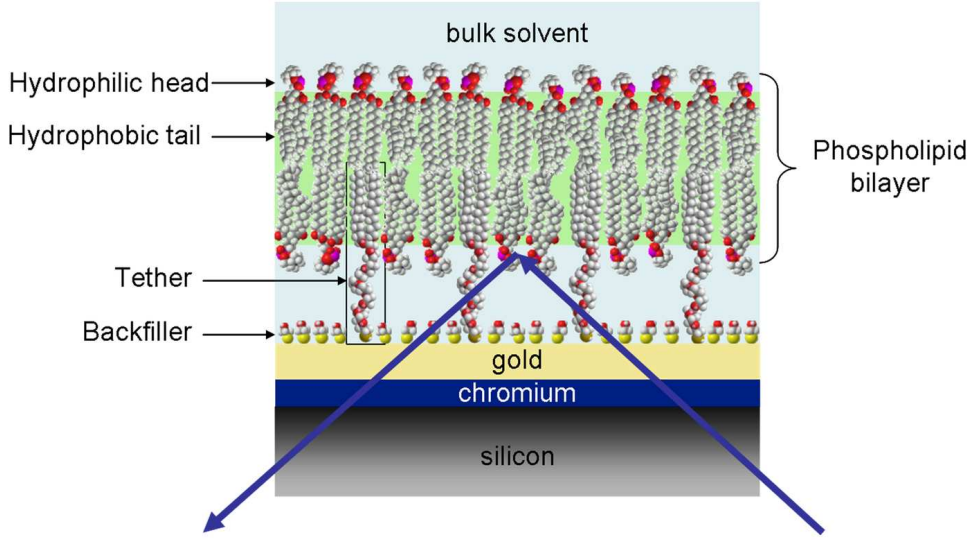


Fig. 7: *Sparsely tethered biomimetic membrane. The arrows show the geometry of the neutron reflection experiment: the neutrons hit the sample from the side of the silicon substrate. After [7].*

Neutron reflection is uniquely capable of characterizing in molecular detail the resulting membrane structure, particularly with respect to the thin hydration layer. Fig. 8 shows data sets measured on the system discussed above in three distinct solvent contrasts (pure D_2O, pure H_2O and a H_2O/D_2O mixture with SLD 4×10^{-6} Å^{-2} called CM4). The three measurements used the same sample, with the exchange of the isotopically distinct water solutions performed *in-situ*, and were co-refined on the assumption that the sample is identical except for the distinct solvent contrasts. As is evident from the differences in the SLD profiles, the spacer region containing the tethers is highly hydrated while the bilayer membrane covers the substrate homogeneously: only a minimal hydration of the inner leaflet is present. Such a system thus constitutes a biomimetic membrane that is very well suited to understand the interaction between lipid bilayer and membrane proteins [29, 30].
The top graph of Fig. 8 shows measurements performed on a system with a molar proportion of tether to backfiller at the gold surface equal to 30 to 70 and the bottom graph shows measurements performed with this proportion equal to 15 to 85. The SLD profiles show that the

aqueous reservoir for the system containing 30 % of tether at the gold surface is 21 Å thick, the one for the system containing 15 % of tether being slighlty thinner, i.e. 19 Å thick. This slight difference in thicknesses is revealed by slight shifts in the positions of the intensity maxima and minima in the region of scattering wave vectors around $Q_z = 0.2$ Å^{-1}. This comparison gives an idea of the high resolution of the neutron reflection technique.

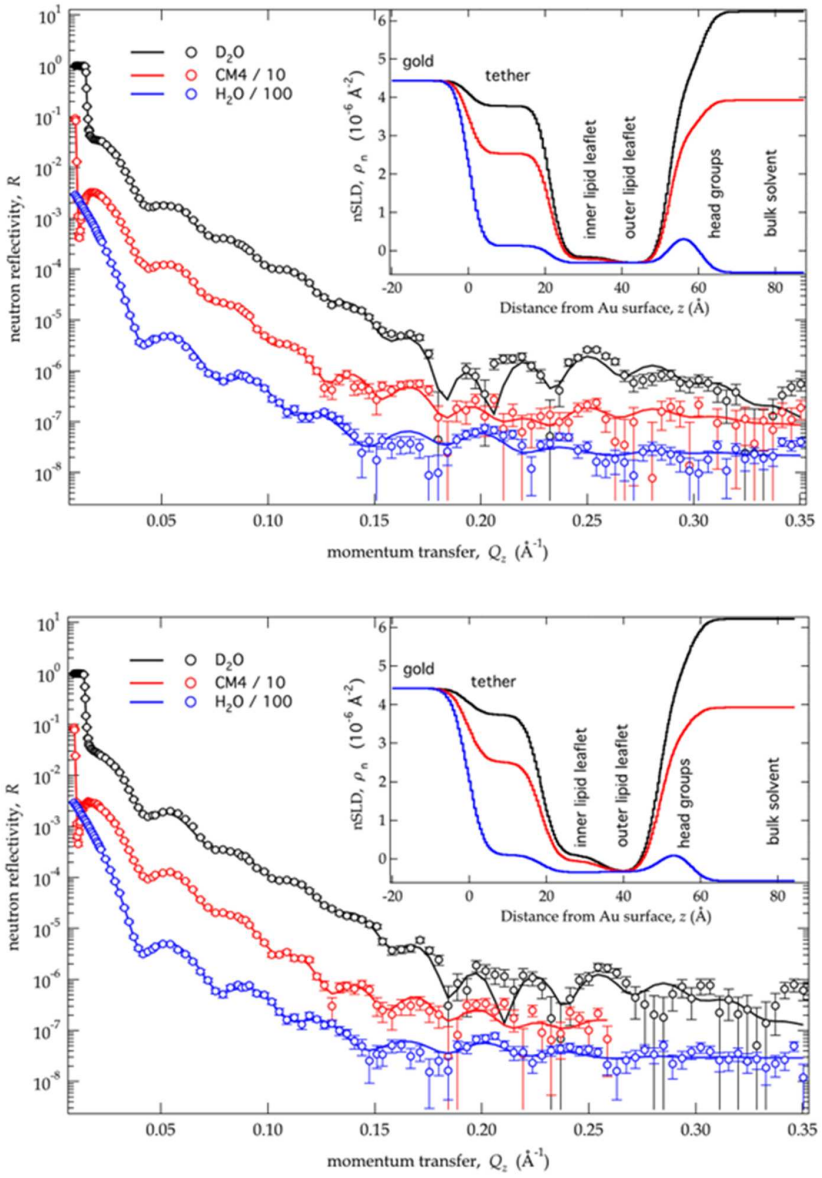


Fig. 8: *Neutron reflectometry on sparsely tethered bilayer lipid membranes. The two graphs represent measurements performed on two different samples with two different molar proportions of tether and backfiller molecules anchored at the gold surface. Top graph: 30 % of tether. Bottom graph: 15 % of tether. After [7].*

4 Exchange bias of CoFe/IrMn heterostructures

4.1 What is exchange bias?

Exchange bias arises in heterostructures consisting of e.g. ferromagnetic and antiferromagnetic thin films [31–33]. After annealing above the Néel temperature of the antiferromagnet (AFM) the sample is cooled down in an external magnetic field H_{FC}. This results in a shift of the hysteresis loop, the so-called exchange bias field H_{EB}, Fig 9. The effect was first observed in 1956 by Meiklejohn and Bean in fine Co particles which had been partially oxidized to CoO which is antiferromagnetic [34].

Though intensely studied, still open questions remain, especially about the microscopic origin of the exchange bias [35]. As a matter of fact, the amplitude of the shift is measured two orders of magnitude smaller than one would expect by taking the hypothesis that the orientations of the atomic magnetic moments are not modified by the interface. As origin for the exchange bias so-called uncompensated interfacial spins are considered that are at the interface of the ferromagnetic/antiferromagnetic film and may result from domains in the antiferromagnet [36–38]. Simple cartoons of the spin configuration within an FM/AFM bilayer are given in Fig. 9. Recently, exchange bias has also been observed in structures nominally not containing any antiferromagnetic material, such as inhomogeneous nanoparticles [39] or transition metal oxide heterostructures [40].

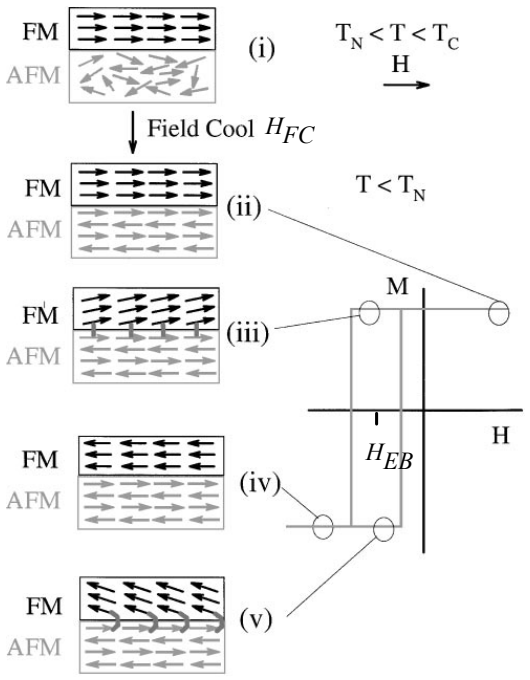


Fig. 9: *Schematic diagram of the spin configuration of an FM–AFM bilayer (a) at different stages (i)–(v) of an exchange biased hysteresis loop (b). Note that the spin configurations are just a simple cartoon to illustrate the effect of the coupling. They are not necessarily accurate portraits of the actual rotation of the FM or AFM magnetizations. Adopted from [31].*

4.2 Sample for the exchange bias effect

To show the power of polarized neutron scattering for investigation of exchange biased systems a $Co_{80}Fe_{20}/Ir_{20}Mn_{80}$ thin film heterostructure has been chosen[41, 42]. The composition of the sample is sketched in Fig. 10. By dc-magnetron sputtering the films were grown on an oxidized Si substrate.

This method is known to produce polycrystalline films. First a NiFe buffer layer is grown on an oxidized Si substrate. The buffer

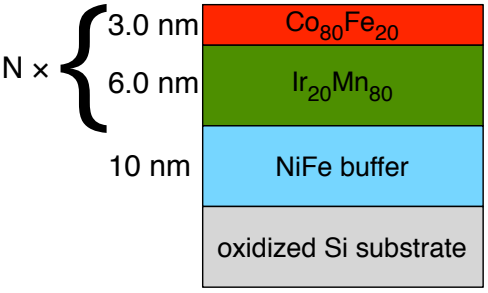


Fig. 10: *Sketch of the sample for demonstrating the exchange bias effect. The number of CoFe/IrMn bilayers is varied from $N = 1$ to $N = 10$.*

layer improves the texture, i.e. the preferred orientation of the crystallites in the subsequent layers. Then, bilayers of 3.0 nm $Co_{80}Fe_{20}$/ 6.0 nm $Ir_{20}Mn_{80}$ are grown. The number of bilayers is varied from $N = 1$ to $N = 10$. While CoFe is a ferromagnet, IrMn is an antiferromagnet with a Néel-Temperature of $T_N = 520$ K. Hence, by heating the sample for 60 min to 533 K, i. e. above the Néel-Temperature of IrMn and then cooling it down to room temperature in an external field $H_{FC} = +130$ Oe a unidirectional anisotropy is induced.

4.3 Magnetic hysteresis measurements

Hysteresis loops of the samples were taken by SQUID (Superconducting Quantum Interference Device) [43], which is an integrating method, Fig. 11. Therefore the loops reflect the magnetic behavior of the heterostructure in total. The shape of the hysteresis implies that there are two different loops, which is most obvious for $N = 1$ in Fig. 11(a). The first loop with low coercivity can be attributed to the hysteresis of the soft magnetic NiFe buffer. The second one with higher coercivity relates to the hysteresis of the CoFe/IrMn heterostructure and exhibits, as expected, a shift of the hysteresis loop to negative fields. The contribution of the FM/AFM heterostructure to the hysteresis loop increases with the number of bilayers. For $N = 10$ the contribution of the NiFe layer is hardly visible. The bias field increases with the number of bilayers from 330 Oe for $N = 1$ to about 900 Oe for $N = 10$ and the shape of the loop changes from rectangular to slanted. How can these observations be explained? To clarify this, structural characterizations have been performed.

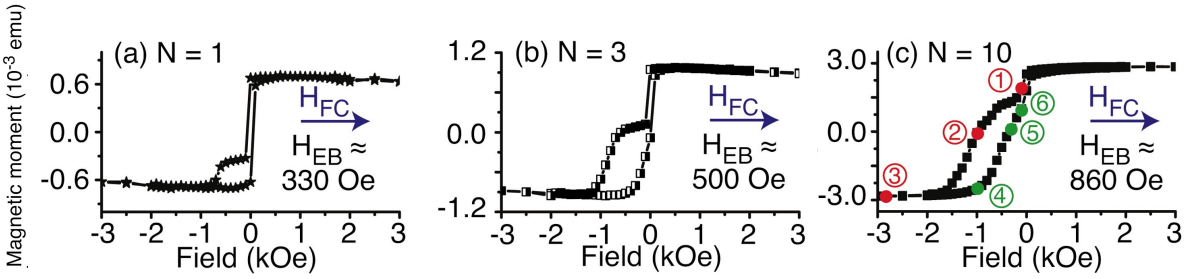


Fig. 11: *SQUID magnetization loops of [3.0 nm $Co_{80}Fe_{20}$/ 6.0 nm $Ir_{20}Mn_{80}$]$_N$/ 10.0 nm NiFe/SiO_2 heterostructures with different numbers of bilayers N. The relative contribution of the NiFe buffer decreases and H_{EB} increases with N. Circled numbers in (c) refer to Fig. 14. Adopted from [41].*

4.4 Structural characterization

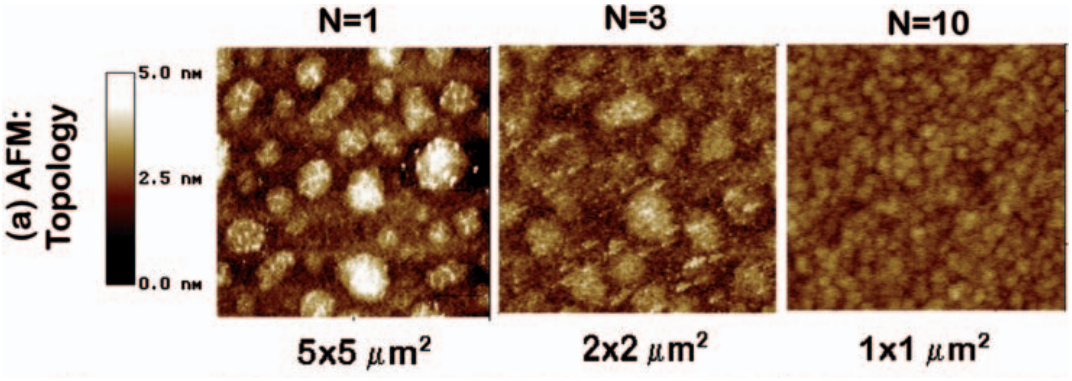


Fig. 12: *Atomic force microscopy micrographs of [3.0 nm $Co_{80}Fe_{20}$/ 6.0 nm $Ir_{20}Mn_{80}]_N$/ 10.0 nm NiFe/SiO_2 heterostructure. Adopted from [42].*

Surface roughness is considered in general to be a possible origin for exchange bias as it may increase the number of uncompensated pinned spins [32]. X-ray scattering investigations on the samples, however, reveal that the surface roughness increases only slightly with the number of bilayers and cannot be regarded as the origin for the strong increase of the bias field [41, 42]. Atomic force microscopy shows a decrease of the grain size with increasing number of bilayers from 650 nm for $N = 1$ to 60 nm for $N = 10$, see Fig. 12. The grain size and the increase of the exchange bias have been plotted in one diagram in Fig 13. The enhancement of the exchange bias field can be correlated with the shrinking of the grain size which is explained in the framework of the domain-state model of the exchange bias [35, 36]. Grain boundaries are energetically preferred sites for domain walls in antiferromagnetic materials [37]. Hence, smaller grains increase the density of domains in the antiferromagnet which leads to an increase of excess magnetization in the antiferromagnet and increases H_{EB} [36]. For multilayers with N of the order of 10 a gradual variation of the grain size and H_{EB} along the growth direction and hence, a nontrivial magnetic reversal behaviour is expected. This will be studied in the next section.

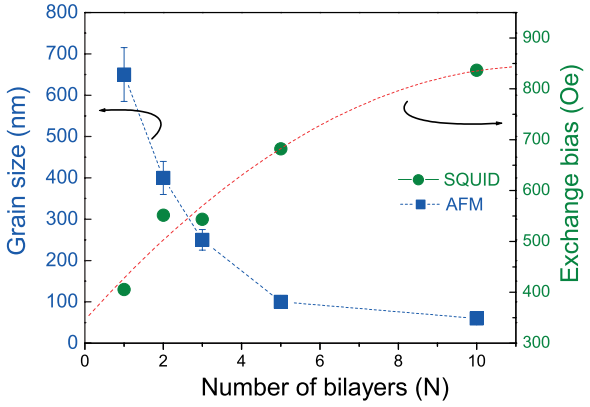


Fig. 13: *Correlation of the dependencies of the grain size and H_{EB} on N. The dashed lines are guides to the eyes. Taken from [42].*

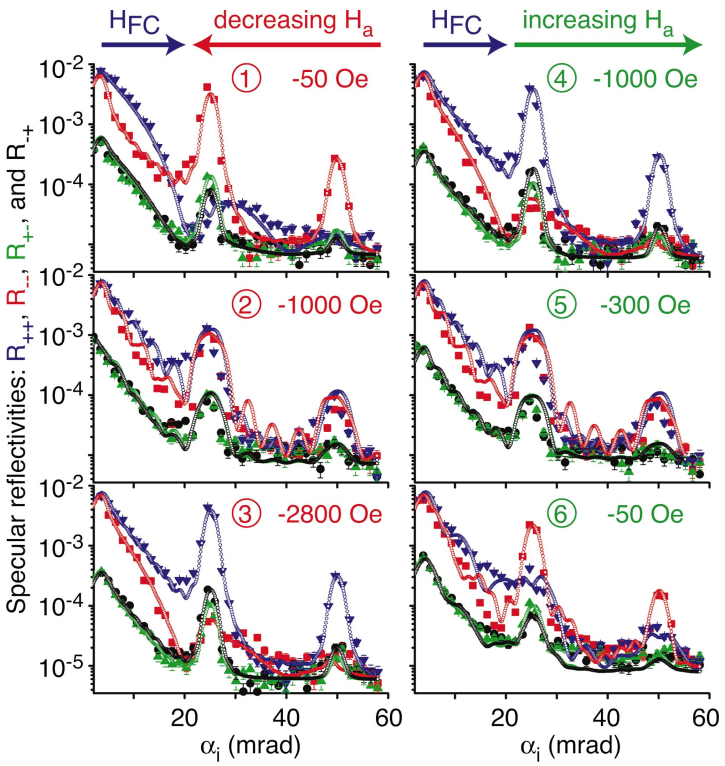


Fig. 14: *Specular reflectivities R_{++} (blue), R_{--} (red), R_{+-} (green), and R_{-+} (black) of a [3.0 nm CoFe/ 6.0 nm IrMn]$_{10}$/ 10.0 nm NiFe/SiO_2 at different positions along the magnetization loop (circled numbers refer to Fig. 11). Filled symbols are the data and open symbols the fits. Taken from [41].*

4.5 Learning about the reversal process of the heterostructure by Polarized Neutron Reflectivity (PNR)

In order to study the magnetic reversal of the CoFe/IrMn heterostructure in dependence of the film depth, that means to resolve the magnetization states of the different layers, polarized neutron reflectivity (PNR) measurements with polarization analysis were performed in specular geometry at the reflectometer HADAS of FZ Jülich[1]. The principle of this measurement was introduced in contribution D2 of this springschool [23]. With PNR four characteristic reflectivities are measured as a function of the incident angle. Spin flip (SF) and non spin flip (NSF) reflectivities are distinguished with respect to the neutron polarization parallel (+) or antiparallel (−) to the external field. When the in-plane magnetic induction of the sample is parallel to the neutron polarization (i.e. the applied field direction) the neutron polarization is not changed and a NSF reflectivity is observed. In contrast, components of the magnetization orthogonal to the neutron spin can flip the polarization of the incident neutron which results in a non-zero SF reflectivity.

Generally, two reversal processes are distinguished. In case of coherent magnetization rotation the magnetization of the sample may be considered as a single magnetization vector which rotates within the sample plane from one field parallel orientation to the opposite. Hence, there

[1]The Jülich research reactor FRJ–2 (DIDO) was shut down in 2006. The successive instrument to the reflectometer HADAS is MARIA (MAgnetic Reflectometer with high Incident Angle) at the FRM II in Garching.

will be a significant increase of the specular reflectivities in the SF channels (R_{+-} and R_{-+}) which are exclusively of magnetic origin and correspond to magnetization components perpendicular to the external field which is also the guiding field for the neutron polarization. Reversal by domain nucleation and propagation does not provide enhanced SF intensities because the magnetization is always collinear to the external field and solely shows up in the specular NSF reflectivities.

PNR measurements of the [3.0 nm $Co_{80}Fe_{20}$/ 6.0 nm $Ir_{20}Mn_{80}]_{10}$/ 10.0 nm NiFe/SiO_2 sample were performed in an external field of different strength. The field direction was collinear to the field used for field cooling, H_{FC}. Fig. 14 shows the specular reflectivities. The two peaks in the NSF channels (R_{++} and R_{--}) at 25 and 50 mrad are the first– and second–order Bragg reflections of the heterostructure. The corresponding weak peaks in the SF channels (R_{+-} and R_{-+}) are due to the limited polarization efficiencies of the setup. The peaks correspond to the bilayer thickness of the sample. The NSF and the SF reflectivities are quite similar in Figs. 14② and 14④ which are taken at fields where the magnetization reversal takes place, see also the hysteresis loop in Fig. 11c). This means, the reversal for both loop branches proceeds via a state with an almost vanishing magnetization component collinear to the external field and a coherent magnetization rotation can be excluded. Moreover, from off-specular measurements Paul et al. even find that the in-plane magnetization component perpendicular to the external field is inhomogenous on a length scale below 1 μm which is rather attributed to fluctuations of this component than to magnetic domains. For all other fields R_{++} or R_{--} have higher intensities and reflect a net magnetization collinear to the external fied, while the specular SF intensities are always much weaker.

The shape of the reflectivities contains further magnetic information which can be retrieved by applying least-square fits based on an extension of the Paratt formalism [44] to magnetic multilayers (open symbols in Fig 14). From off-specular measurements one knows that there is no correlation of the magnetization perpendicular to the film planes [42] which means every layer may be considered to be uncoupled from the others. This information is utilized for the fitting procedure of the specular reflectivities.

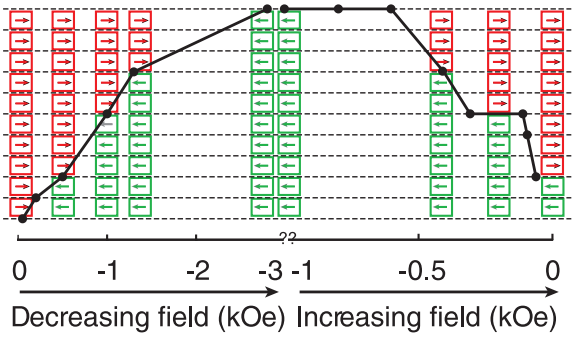


Fig. 15: *Switching sequence along the heterostructure from the bottom to top and back as obtained from the fits. For some field values, each CoFe layer in the heterostructure is represented by a box and the arrows indicate a layer magnetization parallel (red) or antiparallel (green) to H_{FC}. Taken from [42].*

In order to obtain nuclear and magnetic scattering length densities, layer thicknesses and interface roughness, fits of the specular intensities in the saturated states are performed. With these parameters fixed the reflectivities of the non-saturated states are fitted assuming different magnetization configurations, with the magnetization within each layer aligned either parallel or antiparallel to the external field. Decreasing the external field the CoFe layers switch sequentially from the bottom to the top, and increasing the external field the reversal proceeds in the opposite direction, see Fig. 15.

Both, the increasing exchange bias field in the hysteresis loops and the sequential switching of the layers are clearly related to the microstructure, i. e. the decrease of the grain size with increasing number of layers and from bottom to the top within the same multilayer and could be interpreted thanks to polarized neutron scattering.

5 Orbital ordering in transition metal oxide heterostructures

Transition metal oxides are discussed as new materials for electronics [45–47]. One of their most interesting properties is the so-called orbital ordering [48–52]. An overview is given by e. g. Tokura and Nagaosa [17] and Hotta [53].

The d-electrons of a transition-metal ion in an oxide material experience the crystal field of the neighboring atoms which is caused by the electrostatic contributions due to charge transfer and chemical bonding to neighboring atoms. The crystal field theory was discussed in contribution B5 of this springschool. In a free atom or ion the five d-states are degenerate in energy. However, in a crystal, e. g. with perovskite structure ABO_3 (see B5 of this springschool and [54]), the transition-metal ion B is octahedrally coordinated by six oxygen ions O^{2-} which quenches the orbital angular momentum by introducing the crystal field splitting of the d orbitals. Wave functions pointing toward O^{2-} ions have higher energy in comparison with those pointing between them. Hence, there will be a lower lying triply degenerate t_{2g} level and a higher-lying doubly degenerate e_g level. The five d orbitals are depicted in Fig. 16. If additionally, tensile strain acts on the crystal structure, the degeneracy of the energy levels may further be lifted. Hence, the occupation of the d-orbitals is controlled by the crystal structure and strain. Vice versa it controls anisotropic electronic and magnetic properties of many transition metal oxides. [17].

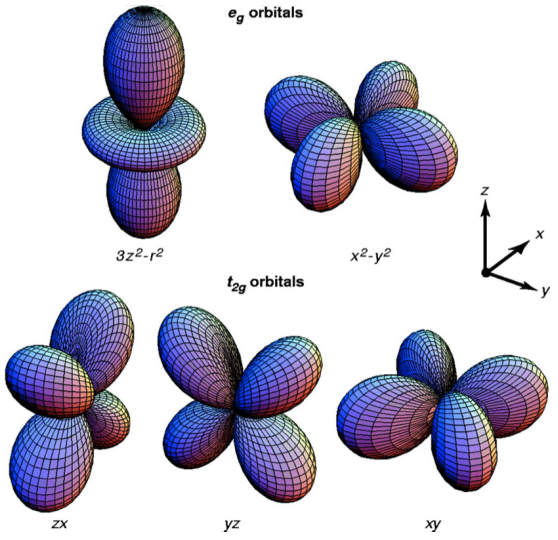


Fig. 16: *Five d-orbitals. In the cubic crystal field the fivefold degeneracy is lifted to two higher-lying e_g orbitals and three lower-lying t_{2g} orbitals. Taken from [17].*

In 2011, Benckiser et al. published their results about an oxide heterostructure consisting of a superlattice of $LaAlO_3$/$LaNiO_3$ on a $SrTiO_3$ substrate [55] which will be discussed in the following. This paper combines scattering and X-ray absorption spectroscopy to gain comprehensive insight into the sample's properties. Within the framework of this lecture only parts of the paper are discussed.

For $LaAlO_3$ (LAO) the electronic configuration is $La^{3+}Al^{3+}O_3^{2-}$. That means, there are no free electrons and LAO is a band isolator. The situation is different for $LaNiO_3$ (LNO): $La^{3+}Ni^{3+}O_3^{2-}$. In this case, there are 7 electrons left, the Ni cation has the $3d^7$ configuration. So the question is: How do the remaining seven electrons occupy the five orbitals? Due to the cubic crystal field a splitting in e_g and t_{2g} levels may be expected. Further lifting of the levels may be expected, if the LAO/LNO heterostructure is exposed to additional strain as for epitaxial growth of the films on a STO substrate. Epitaxial growth means that the in-plane lattice of the substrate is continued by the thin film grown on it. Model calculations for this system suggest, that the degeneracy of the Ni bulk e_g orbitals in LNO is lifted and that the $x^2 - y^2$ level is preferred [55]. This can be understood in terms that tensile strain increases the distance to the in-plane oxygen ions while the distance to the ones in vertical direction is decreased, that is the direction of the $3z^2 - r^2$-orbital. To summarize, Benckiser et al. [55] evidence this prediction experimentally.

5.1 The transition metal heterostructure sample

A [4 u.c. LAO/ 4 u.c. LNO]$_8$ heterostructure was grown on a single-crystalline [001]–oriented STO substrate by pulsed-laser deposition [55], see Fig. 17 (u.c., pseudo-cubic unit cell). X-ray diffraction measurements confirm tensile strain on the LAO/LNO multilayered film due to the STO substrate while the LAO/LNO multilayered film may be considered with averaged lattice constants.

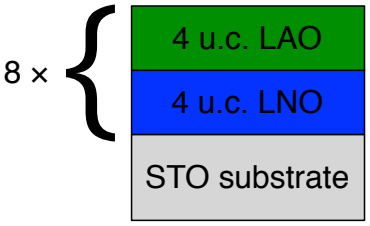
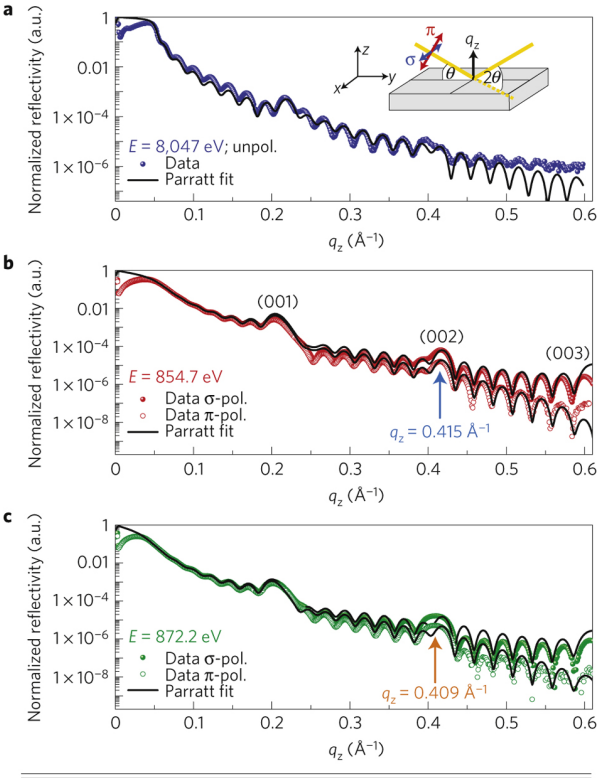


Fig. 17: *Sketch of the LNO-LAO heterostructure on the STO substrate*

5.2 X-ray reflectivity and X-ray absorption spectroscopy



Fit parameters	Thickness (Å)	Roughness (Å)
7 bilayer LAO/LNO (coupled)	15.2/15.4	3.5/2.0
STO substrate	∞	1.7
LNO buffer layer	18.5	3.3
LAO cap layer	18.3	3.0

Fig. 18: *Momentum-dependent X-ray reflectivity of the [4 u.c LAO/ 4 u.c. LNO]$_8$ superlattice. (a) Reflectivity for E = 8047 eV (Cu K_α; hard X-rays), (b) E = 854.7 eV (Ni L_3) and (c) E = 872.2 eV (Ni L_2). All data have been normalized to unity at $q_z = 0$. The measurements in the soft-X-ray range have been made with σ- and π-polarization of the incident X-rays (see the sketch in a). The solid black line is the best fit to the data when using the Parratt algorithm. The fitted parameters, roughness and thickness, are summarized in the table below. The estimated error bars of the individual thicknesses and roughnesses are approximately ± 0.5 Å. Taken from [55].*

X-ray reflectivity measurements were performed at room temperature in specular geometry, that is, the momentum transfer q was parallel to the surface normal z. In Fig.18a, the resulting data for a hard X-ray energy of 8047 eV, which is far from resonance, is shown. The data reflect the high quality of the investigated superstructure. At energies corresponding to the Ni $L_{2,3}$ absorption edges, i. e. 854.7 eV and 872.2 eV, the reflectivity was obtained utilizing linearly σ- and π-polarized light, Figs.18b and c. Clearly, a dependence of the reflectivities on

the polarization of the light can be seen which shows that the X-ray absorption is polarization dependent. Superlattice peaks up to the third order, denoted by SL $(00l)$, $l = 1, 2, 3$ can be observed. The layer thickness and interface roughness of the superlattice were obtained by the authors [55] utilizing a fit program which is based on the Parrat formalism [44]. The resulting profiles are shown in Figs. 18a-c and the structural parameters are given in the table below the figures. It turns out that different interface roughness and thickness has to be considered for the first LNO layer on STO (acting as buffer layer) and the last LAO layer stack (acting as cap layer) for the best fit.

To gain insight into the orbital occupation X-ray absorption spectra (XAS) were measured in total-electron yield (TEY) and fluorescence-yield (FY) modes, Fig. 19. The TEY is interface sensitive while the FY is bulk sensitive. As both measurement modes reveal nearly identical results, only the fluorescence-yield mode is discussed in the following. The principle of this method is introduced in contribution F4 of this springschool. The Ni L_2 and L_3 absorption edges are located around 872 eV and 855 eV, respectively. First, the white line[2] of the Ni L_3 absorption edge has to be separated from the strong white line of La M_4 by subtracting the latter in form of an approximated Lorentzian profile. The result is shown in Fig. 19b. The dichroic difference spectrum in Fig. 19c clearly shows dips at the Ni L_3 and L_2 white line energies, which the authors attribute to natural linear dichroism. Remember, that the different absorption is also visible in the reflectivity data which were taken at the same energies, Fig. 18.

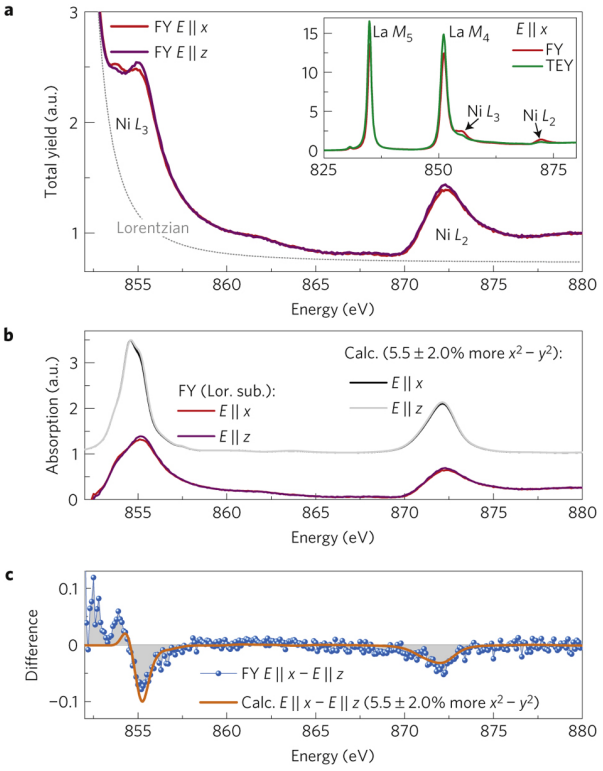


Fig. 19: *Polarization-dependent XAS spectrum (FY) across the Ni $L_{2,3}$ edges. a, XAS spectra for E ∥ x (in-plane) and E ∥ z (out-of-plane) polarization. In the inset Benckiser et al. compare TEY and FY spectra for E ∥ x in the full energy range including the La $M_{4,5}$ white lines. b, Polarized FY spectra after subtraction of a Lorentzian profile fitted to the La M4 line shown together with results for Ni^{3+} XAS spectra with 5.5% higher $x^2 - y^2$ occupation, obtained from the cluster calculation. c, Difference spectra (E ∥ x-E ∥ z) calculated from the measured (blue points) and calculated (orange line) data shown in the middle panel. Taken from [55].*

In general, natural linear dichroism reflects an anisotropy of the charge distribution around a particular ion. To obtain a quantitative estimation of the imbalance in the e_g band occupation in the heterostructure, Benckiser et al. applied the sum rule for linear dichroism [56, 57], which relates the total integrated intensity of the polarized spectra with different polarizations ($I_{E\|x,z}$) to the hole occupation $\bar{n}_{3z^2-r^2}$ and $\bar{n}_{x^2-y^2}$ in the e_g orbitals. After integration in the range of 853 - 877 eV of the spectra which were separated from the La M_4

[2]The name stems from former times when X-ray absorption spectra were taken with the use of photographic plates. The narrow peaks of the absorption edges appeared as unexposed bands on the plate (developed in negative).

contribution, a ratio of the hole occupation of $\frac{\bar{n}_{3z^2-r^2}}{\bar{n}_{x^2-y^2}} = 1.030(5)$ is obtained. From this the orbital polarization

$$P = \frac{n_{x^2-y^2} - n_{3z^2-r^2}}{n_{x^2-y^2} + n_{3z^2-r^2}} = 5 \pm 2\%$$

is derived with $n_{3z^2-r^2}$ and $n_{x^2-y^2}$ the numbers of electrons. Theoretical calculations find an orbital polarization of 5.5 ± 2 % more $n_{x^2-y^2}$ than $n_{3z^2-r^2}$ which means the values agree very well and the slightly higher occupation of the $x^2 - y^2$ orbital is confirmed.

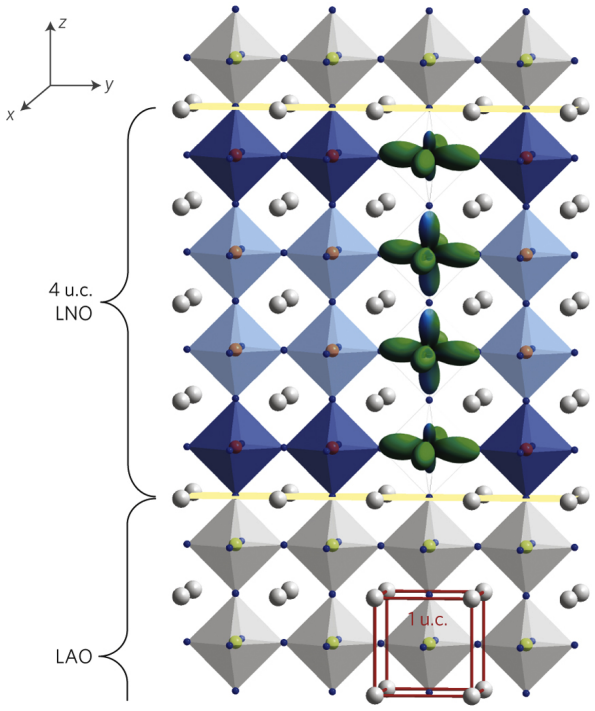


Fig. 20: *Sketch of the LNO/LAO superlattice with layer stacks of four pseudo-cubic unit cells (u.c., see the red box) investigated in this work. The modulation of the Ni 3d e_g orbital occupation along the superlattice normal z is depicted by a different mixture of x^2-y^2 and $3z^2-r^2$ orbitals. The orbital occupation imbalance has been overstated for clarity. Adopted from [55].*

In summary, by combining X-ray scattering and X-ray absorption spectroscopy it was found that the degeneracy of the Ni 3d e_g orbital is lifted along the superlattice normal z in the LNO/LAO superlattice grown on STO. A sketch of the different mixture of $x^2 - y^2$ and $3z^2 - r^2$ orbitals is given in Fig. 20. Further investigation reveals that the population of the different orbitals depends also on the neighboring layer. The discussion of the latter result is far beyond of the scope of this lecture but discussed by Benckiser et al. [55].

6 Conclusion and perspectives

In this chapter, we have shown several investigations of thin film heterostructures from different scientific fields like soft matter, biology, magnetism and electronic oxides. We have concentrated on scattering methods under grazing incidence, because they are best suited to determine the different types of ordering very close to the surface. The interpretation of the data relies a lot on modelling and big efforts are nowadays performed in this direction. A limiting factor for the application of those methods is that they are intensity limited, especially at neutron sources. Optimized instumentation at present high flux reactor sources and at present and future spallation neutron sources are pushing this limit.

References

[1] B.M. Ocko, A. Braslau, P.S. Pershan, J. Als-Nielsen, M. Deutsch, Phys. Rev. Lett. **57**, 94 (1986).

[2] P. Lang, J. Phys.: Condens. Matter **16**, R699 (2004).

[3] Langmuir **25(7)**, (2009)

[4] **Polymer thin films**, Series in Soft Condensed Matter - Vol. 1, edited by Ophelia K.C. Tsui & Thomas P. Russel, World Scientific (2008).

[5] "Nanomaterials for the Life Sciences", Vol.5: "Nanostructured Thin Films and Surfaces", Edited by Challa S. S. R. Kumarlg, Willey (2010)

[6] http://www.ncnr.nist.gov/programs/reflect/rp/biology/index.html

[7] F. Heinrich, T. Ng, D.J. Vanderah, P. Shekhar, M. Mihailescu, H. Nanda, M. Lösche, Langmuir **25**, 4219 (2009).

[8] P. Grünberg, J. Phys. Condens. Matter **13**, 7691 (2001).

[9] D. E. Bürgler, in *36th Spring School 2005, Institute of Solid State Research, Magnetism goes Nano*, (Forschungszentrum Jülich, Series "Mater and Materials", 2005).

[10] J.I. Martin, J. Nogues, K. Liu, J.L. Vincent, I.K. Schuller, J. Magn. Magn. Mater. **256**, 449 (2003).

[11] C.A. Ross, S. Harani, F.J. Castaño, Y. Hao, M. Hwang, M. Shima, J.Y. Cheng, B. Vogeli, M. Farhoud, M. Walsh, H.I. Smith, J. Appl. Phys. **91**, 6848 (2002).

[12] S. Sun, C.B. Murray, D. Weller, L. Folks, A. Moser, Science **287**, 1989 (2000).

[13] J.Y. Cheng, W. Jung, C.A. Ross, Phys. Rev. B **70**, 064417 (2004).

[14] R. Skomskii, J. Phys. Condens. Matter **15**, R841 (2003).

[15] S. Da Col, M. Darques, O. Fruchart, L. Cagnon, Appl. Phys. Lett. **98**, 112501 (2011).

[16] *41st IFF Spring school 2010, Electronic oxides*, (Forschungszentrum Jülich 2010).

[17] Y. Tokura and N. Nagosa, Science **288**, 462 (2000).

[18] S. Thiel, G. Hammerl, A. Schmeh, C.W. Schneider, J. Mannhart, Science **313**, 1942 (2006).

[19] J. Mannhart, D.G. Schlom, Science **327**, 1607 (2010).

[20] A. Ohtomo, H.Y. Hwang, Nature **427**, 423 (2004).

[21] J. Heber, Nature **459**, 28 (2009)

[22] E. Kentzinger, U. Rücker, B. Toperverg, F. Ott, Th. Brückel, Phys. Rev. B **77**, 104435 (2008).

[23] U. Rücker, Lecture D2 of this book.

[24] Y. Murakami, H. Nakao, T. Matsumura, H. Ohsumi, J. Magn. Magn. Mater., **310**, 723 (2007)

[25] J.M. Tonnerre, M. De Santis, S. Grenier, H.C.N. Tolentino, V. Langlais. E. Bomptempi, M. Garcia-Fernadez, U. Staub, Phys. Rev. Lett. **100**, 157202 (2008).

[26] Y. Wakabayashi, H. Sawa, N. Takubo, M. Nakamura, Y. Ogimoto, K. Miyano, J. of Phys.: Conf. Ser. **211**, 012004 (2010).

[27] M. Wolff, U. Scholz, R. Hock, A. Magerl, V. Leiner, H. Zabel, Phys. Rev. Lett. **92**, 255501 (2004).

[28] Nicole Voss, Diplomarbeit, Friedrich-Alexander Universität Erlangen-Nürnberg, 18 August 2007.

[29] D.J. McGillivray, G. Valincius, F. Heinrich, J.W.F. Robertson, D.J. Vandrah, W. Febo-Ayala, I. Ignatjev, M. Lösche, J.J. Kasianowicz, Biophys. Journ. **96:4**, 1547 (2009).

[30] S.A.K. Datta, F. Heinrich, S. Raghunandan, S. Krueger, J.E. Curtis, A. Rein, H. Nanda, J. Mol. Biol. **406** (2011).

[31] J. Nogués and I. K. Schuller, J. Magn. Magn. Mater. **192**(2), 203 (1999), ISSN 0304-8853.

[32] A. Berkowitz and K. Takano, J. Magn. Magn. Mater. **200**(1-3), 552 (1999), ISSN 0304-8853.

[33] R. C. O'Handley, *Modern Magnetic Materials*, New York (Wiley, 2000).

[34] W. H. Meiklejohn and C. P. Bean, Phys. Rev. **102**, 1413 (1956).

[35] P. Miltényi, M. Gierlings, J. Keller, B. Beschoten, G. Güntherodt, U. Nowak, and K. D. Usadel, Phys. Rev. Lett. **84**, 4224 (2000).

[36] U. Nowak, K. D. Usadel, J. Keller, P. Miltényi, B. Beschoten, and G. Güntherodt, Phys. Rev. B **66**, 014430 (2002).

[37] H. Ohldag, A. Scholl, F. Nolting, E. Arenholz, S. Maat, A. T. Young, M. Carey, and J. Stöhr, Phys. Rev. Lett. **91**, 017203 (2003).

[38] F. Nolting, A. Scholl, J. Stöhr, J. W. Seo, J. Fompeyrine, H. Siegwart, J.-P. Locquet, S. Anders, J. Lüning, E. E. Fullerton, M. F. Toney, M. R. Scheinfein, *et al.*, Nature **405**, 767 (2000).

[39] B. Martinez, X. Obradors, L. Balcells, A. Rouanet, and C. Monty, Phys. Rev. Lett. **80**, 181 (1998).

[40] S. J. Zhu, J. Yuan, B. Y. Zhu, F. C. Zhang, B. Xu, L. X. Cao, X. G. Qiu, B. R. Zhao, and P. X. Zhang, Appl. Phys. Lett. **90**(11), 112502 (2007), ISSN 00036951.

[41] A. Paul, E. Kentzinger, U. Rücker, D. E. Bürgler, and P. Grünberg, Phys. Rev. B **70**, 224410 (2004).

[42] A. Paul, E. Kentzinger, U. Rücker, D. E. Bürgler, and P. Grünberg, The European Physical Journal B - Condensed Matter and Complex Systems **45**, 249 (2005), ISSN 1434-6028, 10.1140/epjb/e2005-00056-0.

[43] Y. Xiao, in *41st IFF Springschool 2010, Electronic Oxides* (Forschungszentrum Jülich GmbH, 2010), chap. B3.

[44] L. G. Parratt, Phys. Rev. **95**, 359 (1954).

[45] J. Mannhart and D. G. Schlom, Science **327**, 1607 (2010).

[46] C. Cen, S. Thiel, G. Hammerl, K. E. Schneider, C. W. an Andersen, C. S. Hellberg, J. Mannhart, and J. Levy, Nat. Mater. **7**, 298 (2008).

[47] G. Rijnders and D. H. A. Blank, Nat. Mater. **7**, 270 (2008).

[48] C. Aruta, G. Ghiringhelli, V. Bisogni, L. Braicovich, N. B. Brookes, A. Tebano, and G. Balestrino, Phys. Rev. **B 80**, 014431 (2009).

[49] M. Salluzzo, J. C. Cezar, N. B. Brookes, V. Bisogni, G. M. De Luca, C. Richter, S. Thiel, J. Mannhart, M. Huijben, A. Brinkman, G. Rijnders, and G. Ghiringhelli, Phys. Rev. Lett. **102**, 166804 (2009).

[50] J. Chakhalian, J. W. Freeland, H.-U. Habermeier, G. Cristiani, G. Khaliullin, M. van Veenendaal, and B. Keimer, Science **318**, 1114 (2007).

[51] G. Jackeli and G. Khaliullin, Phys. Rev. Lett. **101**, 216804 (2008).

[52] A. Tebano, C. Aruta, S. Sanna, P. G. Medaglia, G. Balestrino, A. A. Sidorenko, R. De Renzi, G. Ghiringhelli, L. Braicovich, V. Bisogni, and N. B. Brookes, Phys. Rev. Lett. **100**(13), 137401 (2008).

[53] T. Hotta, Rep. Prog. Phys. **69**(7), 2061 (2006).

[54] R. Bruchhaus, in *41st IFF Springschool 2010, Electronic Oxides* (Forschungszentrum Jülich GmbH, 2010).

[55] E. Benckiser, M. W. Haverkort, S. Brück, E. Goering, S. Macke, A. Frañó, X. Yang, O. K. Andersen, G. Cristiani, H.-U. Habermeier, A. V. Boris, I. Zegkinoglou, *et al.*, Nat. Mat. **10**, 180 (2011).

[56] B. T. Thole and G. van der Laan, Phys. Rev. Lett. **70**, 2499 (1993).

[57] G. van der Laan, J. Phys. Soc. Jpn. **63**, 2393 (1994).

E 6 Magnetic Nanoparticles

A. Feoktystov
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

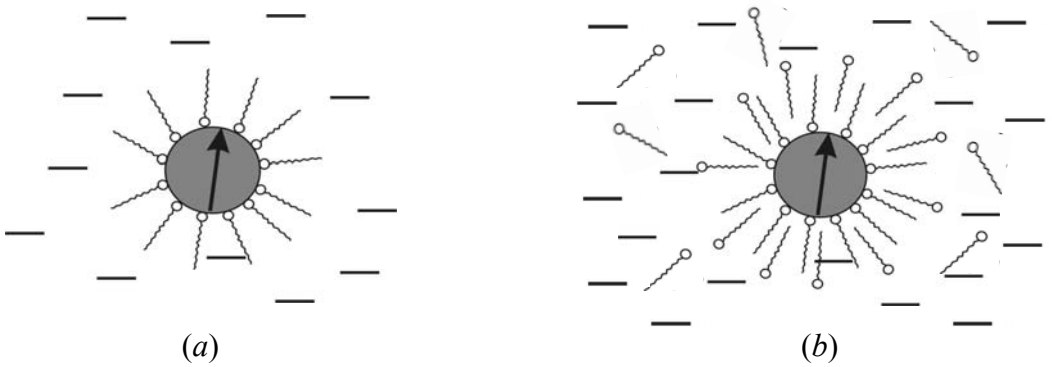
Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Nowadays materials with nanosized particles are of great interest due to their completely different properties as compared to the bulk. The knowledge about the behaviour of such nanomaterials is of great importance. This will give the base for the future development of substances with the given or controlled properties.
In this lecture we will learn quite new and promising substance – magnetic fluid. As a research method, the small-angle neutron scattering (SANS) will be presented. A powerful technique of contrast variation will be discussed in details and applied to small-angle neutron scattering on magnetic fluids. We will see that small-angle neutron scattering investigations of magnetic fluids give important information about particle structure and their interaction.

2 Magnetic Fluids and their Applications

Magnetic fluids (or ferrofluids) are colloidal dispersions of magnetic nanoparticles in liquids. The typical particle size is around 10 nm. The size of the particles is below the critical size, so that such small particles become single domain in contrast to bulk ferromagnetic material which has domain structure of magnetization. The single domain nanoparticles have saturation magnetization. To prevent aggregation of the particles because of the dipole-dipole interaction (especially under influence of the external magnetic field) particles are usually coated with a stabilizing surfactant layer [1]. While in the case of non-polar organic carriers one surfactant layer chemisorbed on the surface of magnetic particles is enough for this purpose (Fig. 1*a*), for polar magnetic fluids the double stabilization, conventionally the formation of the additional second layer, is required [2,3]. In this case, the first layer forms due to chemisorption of surfactant polar heads on the surface of the magnetic particles. The second layer is the result of the physical sorption: tails of the second layer's molecules are turned to the tails of the first layer's molecules, thus polar heads of molecules of the second layer are turned outside, which makes it possible to dissolve the particles in polar carriers. Such kind of double-layer stabilization requires some excess of the surfactant molecules in the solution, so that exchange of molecules of the second stabilization layer with free molecules in the solvent is in equilibrium (Fig. 1*b*).



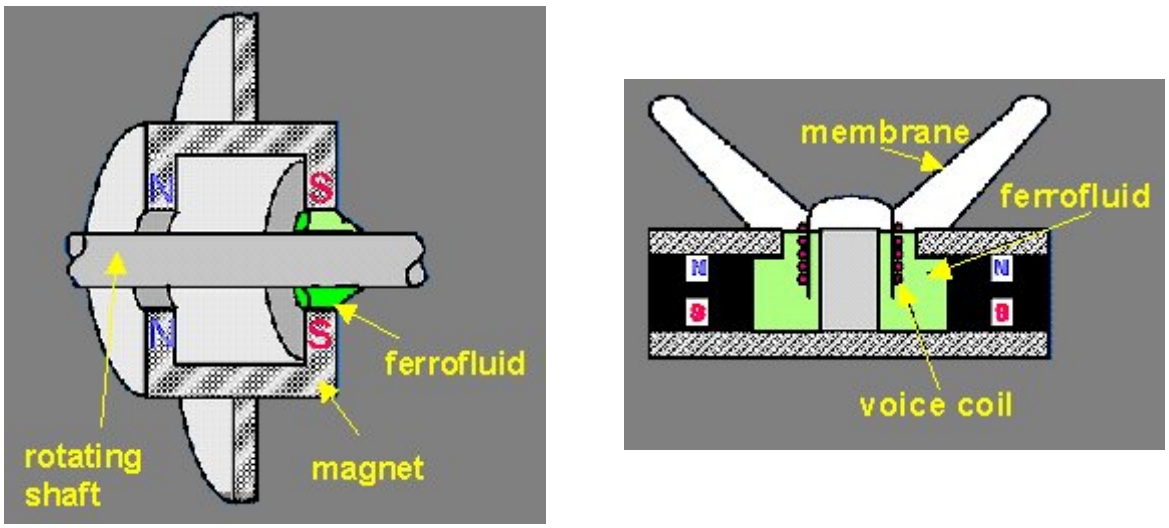
(*a*) (*b*)

Fig. 1: *Schematic representation of nanoparticles in non-polar (a) and polar (b) magnetic fluid. Note excess of surfactant in solution in case of polar ferrofluids.*

Due to the combination of magnetic properties and liquidity ferrofluids found a large variety of applications. All the applications are based on the principles that under the influence of magnetic field the ferrofluid goes to the point where the magnetic field is the strongest, absorption of electromagnetic energy at convenient frequencies causes the ferrofluid heat up and the ferrofluid properties can be changed with the application of magnetic field.

2.1 Technical applications

Technical applications consider the ferrofluids' properties as a whole. In techniques magnetic fluids are already used in dynamic sealing, damping, bearings, separation, heat dissipation, measuring devices and so on [4,5]. We will stop here in more details on the applications of magnetic fluids for heat dissipation and dynamic sealing (Fig. 2).



(*a*) (*b*)

Fig. 2: *Technical application of ferrofluids for vacuum sealing (a) and heat dissipation (b).*

Ferrofluids are used to form liquid seals around the spinning drive shafts in hard disks (Fig. 2*a*), which have to operate in a hermetically closed box because any grain of powder or even smoke may spoil the reading and writing process. This is achieved by making the hole inside a magnet and the shaft made of soft magnetic material. A groove in the shaft is filled up with ferrofluid, which is kept in place by the magnetic field, obstructing the passage of any impurity, but leaving the axle free to rotate, because the obstructing material is liquid [4,5].
Another good example of ferrofluid application is loudspeaker (Fig. 2*b*). Its coil heats up by functioning and the ferrofluid is kept in place by the magnetic field of the magnet which is fixed on the loudspeaker's horn. When the magnetic fluid temperature reaches Curie temperature, ferrofluid's particles lose their magnetic properties and a non-magnetic liquid would flow away and will be substituted by a part of the fluid which has not been overheated. In this case a passive heat transfer is realized. Nowadays most of the high power loudspeakers are equipped with ferrofluid. The presence of the fluid around the coil improves also the

quality of the speaker because it damps unwanted resonances, which would produce a very unpleasant noise [5].

2.2 Biomedical applications

In contrast to technical applications biomedical applications of magnetic fluids focus on the single colloids' properties. The main direction of ferrofluids' application in biomedicine is cancer treatment. It is magnetic hyperthermia [4,5,6], drug targeting and delivery [5,7], contrast medium in Magnetic Resonance Imaging (MRI) [5,6,7].

When one applies a drug onto the surface of magnetic particle and focuses the nanoparticles around cancer tumor with the external magnetic field, the amount of drug necessary is much less than what would be necessary if it were dispersed in the whole body. When the magnetic field is turned off the drug will disperse in the body, but, since the total amount is very small, there will be practically no side effects [5].

The powerful technique of MRI is based on the different relaxation times T_2 of the proton's magnetic moments when it is inside different environments. Often the differences are not strong enough to obtain well resolved images. Dextran coated iron oxides are biocompatible and are excreted via the liver after the treatment. They are selectively taken up by the reticuloendothelial system. This is important because tumor cells do not have the effective reticuloendothelial system of healthy cells, so that their relaxation time is not altered by the contrast agent, which makes them distinguishable from the surrounding healthy cells [5,8].

Hyperthermia is a promising therapy technique which is based on the property of ferrofluids to absorb electromagnetic energy at a certain frequency. This allows one to heat up a localized portion of a living body, where ferrofluid has been injected, for example a tumor, without heating at the same time the surrounding parts of the body. The results of successful experiments healing cancer tumors in rats and rabbits can be found in [9].

The main problem of Hyperthermia is the high Curie temperature of the used ferromagnetic materials. The solution is to define substances with low Curie temperature in order to avoid possible harmful overheating of the human body [10].

3 Magnetic Properties of the Nanoparticles

In bulk ferromagnets the magnetization has a domain structure. These domains form to minimize the magnetostatic energy of the material. As the size of a magnetic particle is reduced down to the critical one the magnetic properties of the material are then dictated by the particle anisotropy and shape rather than by the microstructure of the bulk magnets. There are two main sources of magnetic anisotropy – magnetocrystalline anisotropy and shape anisotropy. Magnetocrytalline anisotropy is determined by the atomic structure of a crystal, which introduces preferential directions for the magnetization (easy axes). The easy and hard directions arise from the interaction of the spin magnetic moment with the crystal lattice (spin-orbit coupling). Shape anisotropy appears when a particle is not perfectly spherical. The demagnetizing field will not be equal for all directions, creating one or more easy axes.

Let's have a look on coercivity as a function of particle size (Fig. 3) [11]. For large sizes the particles are multi-domain becoming more bulk-like with increasing size. Below a critical particle size domain walls will no longer form due to energy considerations and single domain

particles are stable. This critical size corresponds to the peak in the coercivity in Fig. 3. The peak position on the particle size axis is dependent on the values of anisotropy and magnetization. Below the coercivity maximum the particles remain stable until the effects of temperature destroy the ferromagnetic order. The particles are then superparamagnetic. The superparamagnetic size strictly depends on the magnetocrystalline anisotropy of the material [11].

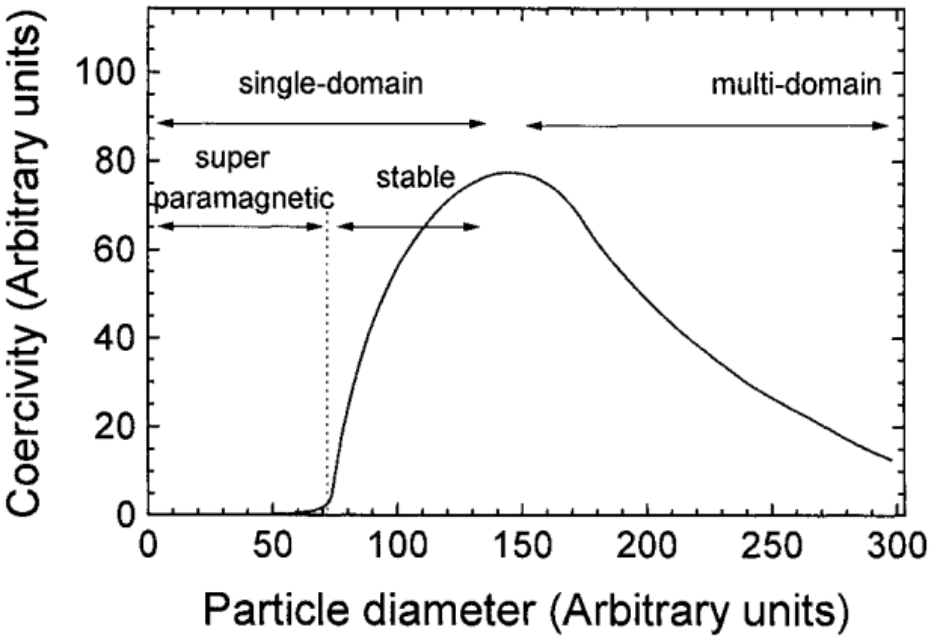


Fig. 3: *Coercivity as a function of particle size. From [11].*

3.1 Superparamagnetism

An assembly of non-interacting single-domain isotropic particles behaves like classical paramagnetic matter but with very high ($\sim 10^3 - 10^4\ \mu_B$) effective magnetic moment μ per particle [12]. If the influence of thermal energy is high enough the magnetic moments of the nanoparticle are randomized unless a magnetic field is applied. In thermal equilibrium the average magnetic moment of assemble has the following expression [12]:

$$< \mu >= \mu \cdot (\coth(\mu H / kT) - kT / \mu H) . \qquad (1)$$

Such substances with a huge magnetic susceptibility, whose average magnetic moment is represented by equation (1), are called then superparamagnetic.

3.2 Néel relaxation. Blocking temperature

Equation (1) is valid only in case of negligible anisotropy energy, but real single-domain particles are anisotropic. In the normal conditions the magnetization direction inside particle is along the easy axis. The two states of magnetization of a uniaxial magnetic particle are separated by an energy barrier (Fig. 4), $K_u V$, where K, is the anisotropy energy density and V

is the particle volume [11]. If the thermal energy, kT, becomes comparable to the barrier height there is an increased probability of the magnetization reversing.

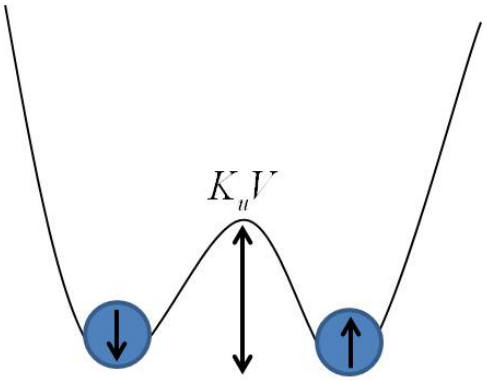


Fig. 4: *Double potential well with two possible orientations of particle magnetic moment.*

This thermally activated switching in case of non-interacting single-domain nanoparticles with the uniaxial magnetic symmetry follow the Née1 relaxation law [12]:

$$\tau = \tau_0 \exp(\frac{K_u V}{kT}), \tag{2}$$

where τ is the relaxation time and τ_0 is a constant.
In case of hard disks stability over a time scale of, for example, 10 years gives a stability criterion of $K_uV/kT > 40$ [13].
The value of τ_0 is typically in the range 10^{-13} –10^{-9} s [12,14]. The actual magnetic behavior of nanoparticle assemble depends on the value of the measuring time, τ_{exp}, of the specific experimental technique with respect to the relaxation time.
If $\tau_{exp} >> \tau$, the relaxation is fast and a time average of the magnetization orientation is observed during the measurement time. In this case the assembly of nanoparticles behaves like a paramagnetic system. If $\tau_{exp} << \tau$, the relaxation is slow and non-equilibrium properties are observed. The two cases are separated by the blocking temperature, T_B. It can be defined from the expression $\tau_{exp} = \tau$:

$$T_B = \frac{K_u V}{k} \ln^{-1}\left(\frac{\tau_{exp}}{\tau_0}\right). \tag{3}$$

Near T_B changes of magnetization reorientation occur with relaxation times comparable with the time of a measurement. Well below T_B the thermal agitation can be neglected and magnetization has static behavior.
It is quite hard to define T_B uniquely. The values of τ_{exp} depend on the experimental technique. In case of Mossbauer spectroscopy τ_{exp} is of the order of a few nanoseconds while in static magnetization measurements the timescale is typically of the order of 1 s. Therefore,

the same nanoparticles can be in the superparamagnetic state in static magnetization experiments and in the blocked state in Mossbauer spectroscopy experiments [12,14].
Nevertheless, there are some general properties of the T_B: the blocking temperature increases with nanoparticle size increase and for a given particle size increases with decreasing measuring time. The highest possible value of T_B is represented by the Curie (or Néel) temperature, when the magnetic moments of each particle disappear [12].

4 Small-Angle Neutron Scattering

There are several possible ways for neutron to interact with a matter. In general, neutrons can be absorbed or captured with all the corresponding consequences or neutrons can scatter on the atoms and inhomogeneities either elastically or inelastically. Small-angle neutron scattering (SANS) is an elastic coherent scattering of neutrons [15]. The typical neutron wavelength used in SANS experiments is in the order of several angstroms [16]. Since distances between atoms in a substance are of the same order as neutron wavelength the coherent scattering picture is concentrated in small-angle region. That is why this method received its name. In SANS experiment one registers a dependence of scattering cross-section (scattering intensity I) on modulus of scattering vector q (i.e. on scattering angle):

$$q = \frac{4\pi}{\lambda}\sin\theta, \tag{4}$$

where λ is the neutron wavelength and 2θ is the scattering angle.
The intensity of small-angle neutron scattering is a superposition of particle form- and structure factors. From the form-factor dependence on the scattering angle one obtains information about distribution of the scattering density inside scattering particles (size, shape, etc.). The structure-factor gives information about particle displacement (radial distribution function).
In 1939 Guinier derived a general expression for the q-dependence of scattering intensity at small angles, which gives number of scattering parameters directly from scattering curve [15]:

$$I(q \sim 0) = I(0)\exp\left(-\frac{R_g^2 q^2}{3}\right). \tag{5}$$

Here $I(0)$ is called zero scattering intensity and contains information about amount of the scattering particles and R_g is a gyration radius, which as in mechanics is directly connected to particle dimensions.

4.1 Contrast variation. Basic functions

In contrast to X-ray scattering the amplitude of neutron scattering depends on atomic number in random way. Thus, for example, it made it possible to detect signal from hydrogen containing substances of the background of heavy atoms such as iron, copper and so on.

If we consider particles placed in the certain solvent, we have to take into account that the scattering intensity will not be proportional just to square of the average particle scattering density but to the square of difference between the average scattering density of the particles and scattering density of the solvent. This difference is called contrast. If the scattering densities of the particle and of the solvent are equal the neutrons doesn't "see" the particles, since the solution represents a homogeneous continuum, and waves scattered into any direction will be extinguished. The neutrons give a brilliant possibility to study complex structures "step-by-step" since one can use hydrogen/deuterium substitution in the solvent or in some part of the particle. It will significantly change scattering contrast of the system revealing scattering features that were buried in case of fully protonated system. The idea and theory of such influence on system contrast was developed by H.B. Stuhrmann and is known as contrast variation [17].
He showed that scattering intensity can be expanded into series of contrast [17]:

$$I(q) = I_s(q) + \Delta\rho I_{cs}(q) + (\Delta\rho)^2 I_c(q), \tag{6}$$

where I_s, I_{cs} and I_c are called basic functions, $\Delta\rho$ is a contrast. I_s gives the residual scattering at zero contrast – scattering on density fluctuations. I_c dominates at high contrasts and represents thus the scattering intensity of the particle shape. I_{cs} is a cross-term.
The parameters of Guinier approximation (5) have the following dependences (see for example [18]):

$$I(0) = nV_c^2(\Delta\rho)^2 \text{ and } R_g^2 = R_c^2 + \alpha/\Delta\rho - \beta/(\Delta\rho), \tag{7}$$

where R_c and V_c are particle gyration radius and corresponding volume, n particle number density, α and β are the parameters describing the relative distribution of the "heavy" and "light" components inside particles [16].
This is only valid for monodisperse non-magnetic particles. In case of polydisperse and magnetic particles, whose scattering density also depends on size, one has to make averaging of equation (6) over the particle size distribution function. Nevertheless, as it was shown in [18] it is still possible to derive the final expression to the same form as (6), but with modified basic functions and modified contrast:

$$I(q) = \tilde{I}_s(q) + \Delta\tilde{\rho}\tilde{I}_{cs}(q) + (\Delta\tilde{\rho})^2 \tilde{I}_c(q). \tag{8}$$

Here $\Delta\tilde{\rho}$ is the modified contrast – the difference between effective average scattering density of the particles $\overline{\rho}_e$ and solvent [18].
Since magnetic scattering doesn't depend on contrast, its contribution enters into the $\tilde{I}_s$ modified basic function. Among all of the modified basic functions, which are now the expressions of classical basic functions, the modified basic function $\tilde{I}_c$ is just the average of the classical shape scattering basic function I_c over the polydispersity function.

The parameters of Guinier approximation have qualitatively different from (see [18] for more details):

$$I(0) = n(\Delta\tilde{\rho})^2 < V_c^2 > + n < (\bar{\rho} - \bar{\rho}_e)^2 V_c^2 > + (2/3) n \rho_m^2 < V_m^2 >, \tag{9a}$$

$$\tilde{R}_g^2 = \left[\frac{< V_c^2 R_c^2 >}{< V_c^2 >} + \frac{A}{\Delta\tilde{\rho}} - \frac{B}{(\Delta\tilde{\rho})^2} \right] \Big/ \left[1 + \frac{D}{(\Delta\tilde{\rho})^2} \right], \tag{9b}$$

where ρ_m and V_m are magnetic scattering density and magnetic volume of the particle, A, B, D are parameters and <…> denote averaging over polydispersity function. Now it is impossible to find a scattering density of the solvent which would zero contrast from all of the particles in solution simultaneously, so there is always residual scattering intensity at the effective match point (where $\Delta\tilde{\rho} = 0$): due to polydispersity (the second summand in equation (9*a*)) and contrast independent magnetic scattering (the third summand in equation (9*a*)).

4.2 Core-shell structure

Magnetic nanoparticles in ferrofluids are of core-shell structure. As an example let's consider a simple case of monodisperse non-magnetic nanoparticles (Fig. 5).

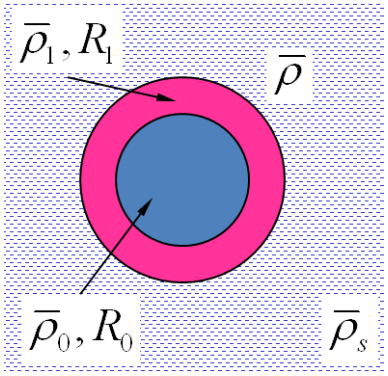


Fig. 5: *Schematic representation of the particle core-shell structure.* $\bar{\rho}_0$ *and* $\bar{\rho}_1$ *are scattering densities of the particle core and shell (*$\bar{\rho} = (V_0/V_1)\rho_0 + (1 - V_0/V_1)\rho_1$ *is the average particle scattering density),* $\bar{\rho}_s$ *is solvent scattering density.* R_0 *and* R_1 *are the radii of particle core and shell.*

The expression for scattering intensity in this case consists of the core and shell contributions:

$$I(q) = n[(\bar{\rho}_1 - \bar{\rho}_s)V_1\Phi(qR_1) - (\bar{\rho}_1 - \bar{\rho}_0)V_0\Phi(qR_0)]^2, \tag{10}$$

where V_0 and V_1 are the particle and "particle+shell" volumes and $\Phi(x)$ is the spherical scattering amplitude: $\Phi(x) = 3(\sin x - x\cos x)/x^3$ [16].
Combining the members at contrast $\Delta\rho = \bar{\rho} - \bar{\rho}_s$, the three basic functions in equation (6) can be derived analytically:

$$I_s(q) = nV_0^2(\overline{\rho}_1 - \overline{\rho}_0)^2[\Phi(qR_1) - \Phi(qR_0)]^2, \quad (11a)$$

$$I_{cs}(q) = 2nV_1V_0(\overline{\rho}_1 - \overline{\rho}_0)\Phi(qR_1)[\Phi(qR_1) - \Phi(qR_0)](\Delta\rho), \quad (11b)$$

$$I_c(q) = nV_1^2\Phi^2(qR_1)(\Delta\rho)^2. \quad (11c)$$

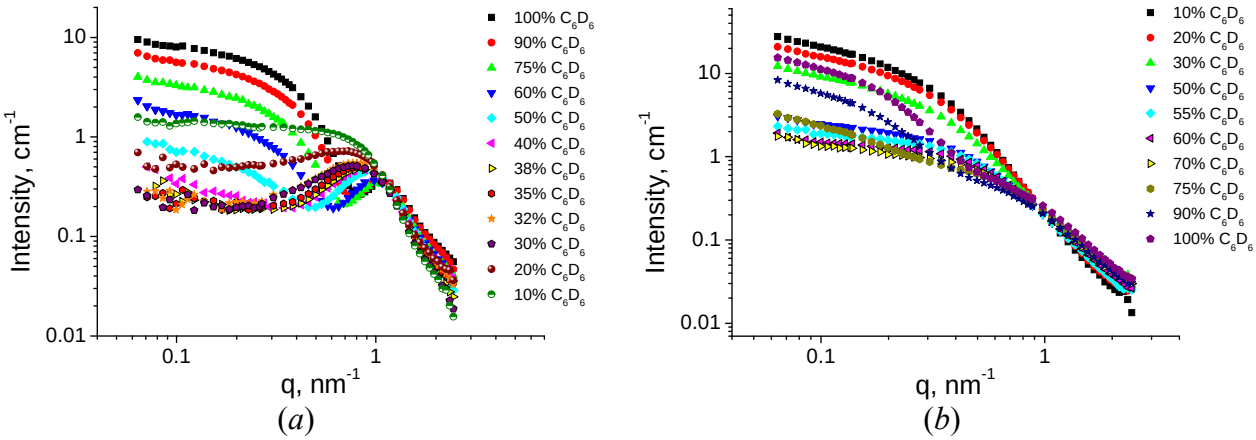
To obtain the three basic functions one has to measure scattering curves for at least three contrasts, then the system of equations of type (6) can be easily solved and compared to the analytical result.
In case of magnetic nanoparticles one has to add a magnetic contribution in equation (10), the square of magnetic form-factor F_M:

$$F_M^2(q) = [\rho_m V_m \Phi(qR_m)]^2. \quad (12)$$

In case of SANS with non-polarized neutrons on diluted unmagnetized samples of ferrofluids, the total scattering intensity contains contributions of nuclear form-factor and 2/3 of magnetic form-factor (for more details refer to [19]):

$$I(q) = F_N^2(q) + \frac{2}{3}F_M^2(q). \quad (13)$$

As a first example we consider organic magnetic fluids which are polydisperse magnetite (Fe_2O_3) nanoparticles coated with a single layer of oleic (OA) or myristic acid (MA) in benzene [20].



(*a*) (*b*)

Fig. 6: *Contrast variation curves of SANS on diluted ferrofluids: (a) – MA stabilization, (b) – OA stabilization. The amount of D-benzene in the solvent (i.e. contrast) is indicated in the legend.*

For changing system contrast initial concentrated magnetic fluids (based on C_6D_6) were diluted with different mixtures of D- and H-benzene, so that the D-benzene content in the final solution was varied within interval 10-100 % (10–12 points). The final magnetic particle

concentrations in ferrofluids were 0.8 % and 0.7 % in case of MA and OA stabilization respectively. The whole set of obtained scattering curves at different contrast are presented in Fig. 6.

To each of the scattering curves in Fig. 6 the Guinier approximation (5) was applied and the dependences (9) were obtained (Fig. 7):

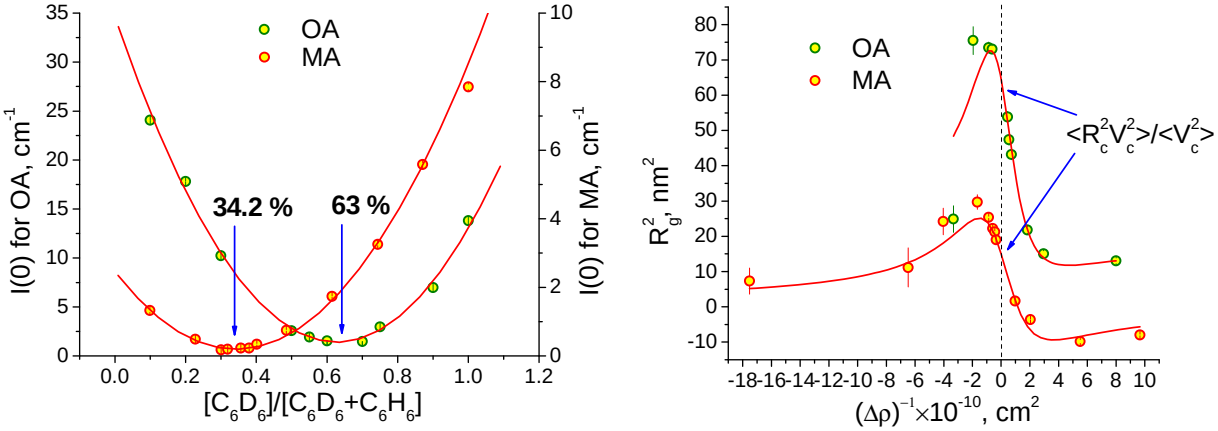


Fig. 7: *The parameters of the Guinier approximation (5) fitted according to equations (9). (a) – zero scattering intensity dependence on deuteration of the solvent and (b) – dependence of gyration radius on inversed contrast.*

The effective match points were determined from parabolic fitting of curves in Fig. 7*a* as 34.2 ± 0.4 % and 63 ± 3 % of the D-benzene content for MA and OA fluids, respectively. The corresponding scattering density's values are (2.64 ± 0.03) × 10^{10} cm^{-2} and (3.86 ± 0.15) × 10^{10} cm^{-2}. Taking into account the quasi-spherical shape of particles in magnetic fluids the characteristic radius of the whole particle was calculated from the fit results were calculated to be 5.0 ± 0.1 nm (MA) and 10.3 ± 0.1 nm (OA) [20]. Thus, combining the results of the calculated effective match point, which shows that OA sample contains more of heavier component (magnetite) in the particles, and radius of the particles one can speak about size regulation effect when using such two surfactants to synthesize the ferrofluids. The particle size, which comprises both the mean size and polydispersity, in OA fluid is twice larger as compared to MA fluid.

From the whole set of scattering curves the three basic functions from equation (8) were calculated. The best way to check the reliability of the calculated basic functions is to compare I_s basic function, which is a residual scattering intensity at zero contrast, with the scattering curve close to math point (Fig. 8). As one can see, the I_s basic functions coincide well with measured SANS curves which confirms correctness if the basic function separation. Comparison of the shape scattering basic functions I_c for two magnetic fluids directly confirms the results of fit of gyration radius dependence on inversed contrast in Fig. 7*b* with the expression (9*b*). The I_c basic function in case of OA stabilization is more steep than MA sample, which is the results of larger particles in the ferrofluid.

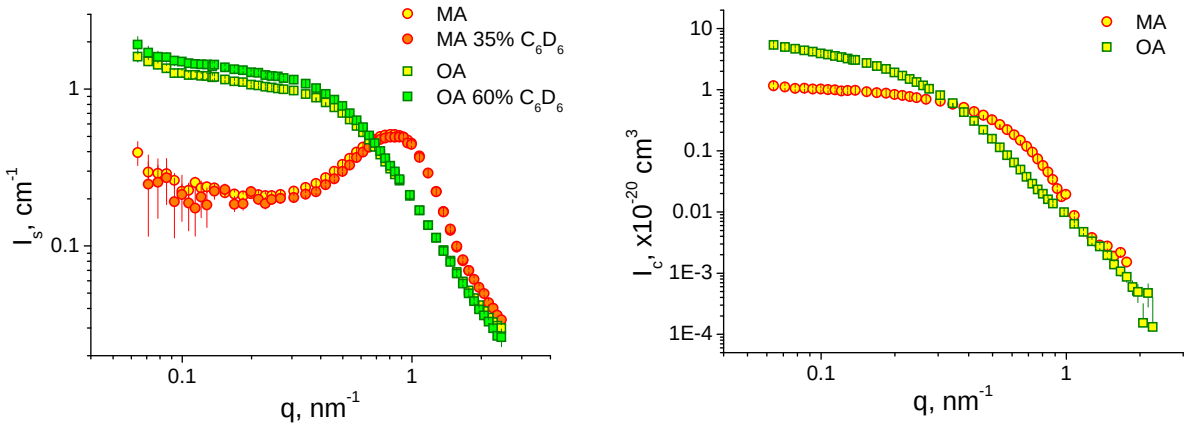


Fig. 8: *Calculated basic functions of expression (8). (a) represents I_s basic function of the residual scattering and (b) – I_c shape scattering basic functions. I_{cs} is not shown due to the lack of its transparent interpretation.*

For the next example we consider water-based magnetic fluids. Performing the same steps for sample dilution with different ratios of H_2O and D_2O as for previous ferrofluids the SANS contrast variation curves were measured in 10 contrast points for each sample (Fig. 9) [21,22].

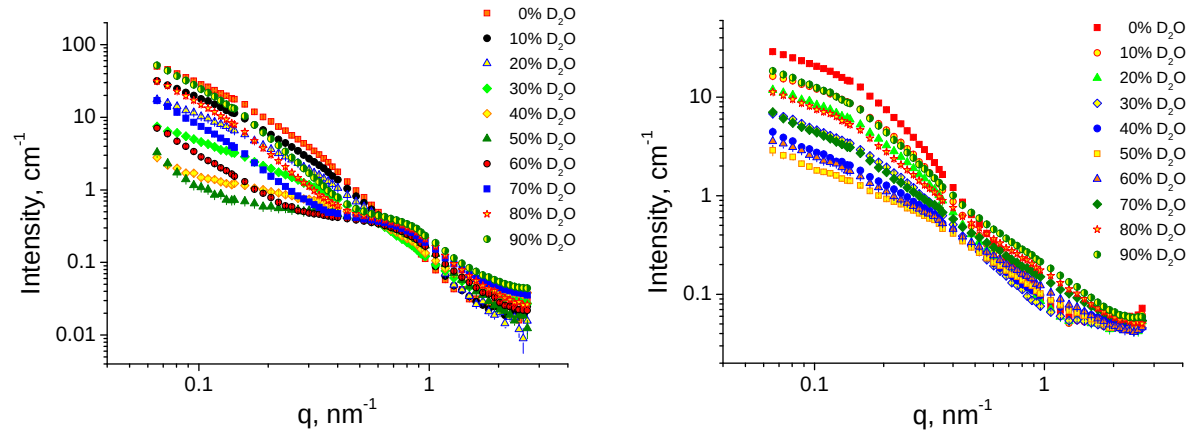
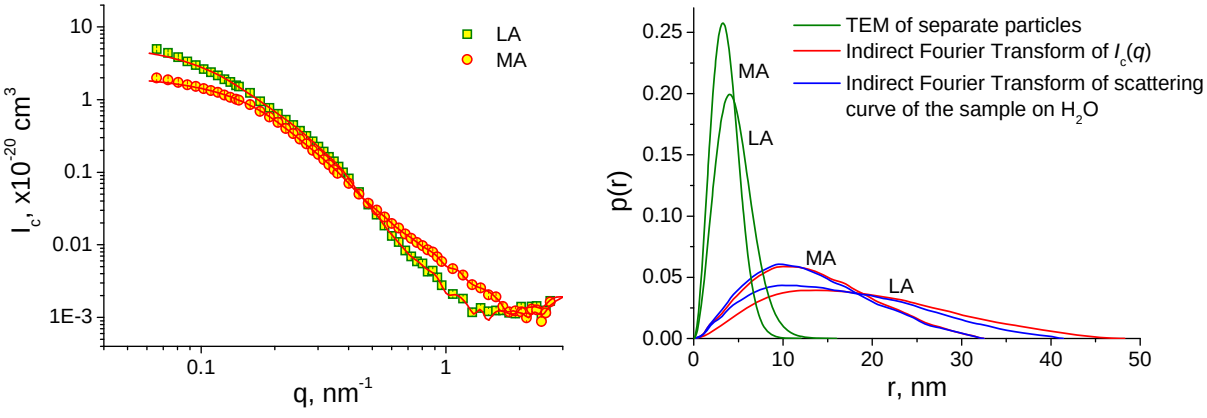


Fig. 9: *SANS contrast variation curves for water-based magnetic fluids with double layers of lauric (LA) (a) and myristic acid (MA) stabilization (b). The ration of D_2O in the solvent is indicated in the legend.*

As one can see, Guinier region, which is represented by a plateau at small *q*-values, is inaccessible here due to an aggregation in ferrofluids. The water-based magnetic fluids require double stabilizing shell around magnetic nanoparticles. The second layer is quite weak and its molecules are in equilibrium with the free surfactant in the solvent. The equilibrium is very sensitive to external actions and such ferrofluids are usually characterized by the existence of aggregates (see for example [23]).

Nevertheless, the three basic functions (8) were calculated, among them the most informative shape scattering basic function I_c (Fig. 10*a*). From the more steep I_c curve of LA magnetic fluid one can easily say that particles (aggregates) in the sample are large than in MA sample.



(*a*) (*b*)

Fig. 10: *Shape scattering basic functions for both ferrofluids. Line shows fit according to Indirect Fourier Transformation [24] (a). The pair distance distribution functions calculated from the shape scattering basic functions, scattering curves of the samples in pure H_2O and compared to the results of Transmission Electron Microscopy measurements [22] (b).*

The comparison of pair distribution functions calculated from I_c basic function and scattering curves of ferrofluids on H_2O, where due to a weak contrast between shell and water the scattering originates mainly from magnetite core, are compared to the results of Transmission Electron Microscopy, when only single particles were taken into account (Fig. 10*b*). In both cases experimental curves give broader $p(r)$ function pointing to the existence of large aggregates in the samples. In case of LA sample the maximum r-value of $p(r)$ functions are different on 7 nm. This can be associated with shell (~3.5 nm) around the particles which is apparent in I_c function, but not in scattering on H_2O-based sample. In MA sample the $p(r)$ functions coincide, so in this sample the aggregates consist of nanomagnetite with one incomplete layer of the surfactant [21,22].

In case of maghemite water-based ferrofluid stabilized by short citric ions (charge stabilization [25]) the particles can be considered as homogeneous [26]. The Coulomb interaction is compensated with the addition of NaCl to the solvent. In that case, the result of SANS contrast variation experiment gives directly two characteristic sizes of the particles – nuclear and magnetic [18]. The dependence of the gyration radius on contrast together with its fit is presented in Fig. 11.

The difference in magnetic and nuclear radii of around 3 nm could hardly be explained as the existence of magnetically dead layer on the edge of maghemite nanoparticles in the ferrofluid. As it was shown such a difference can be explained as a result of non-uniform charge distribution on the surface of the particles and as a result the existence of residual Van der Waals interaction (for more details refer to [26]).

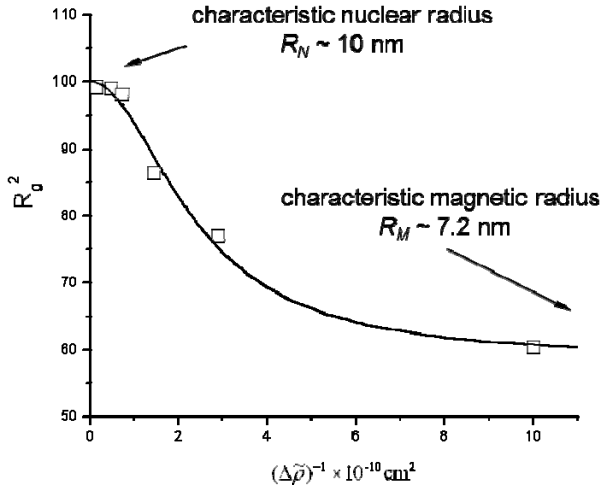


Fig. 11: *The dependence of the gyration radius on contrast and its fit according to [18].*

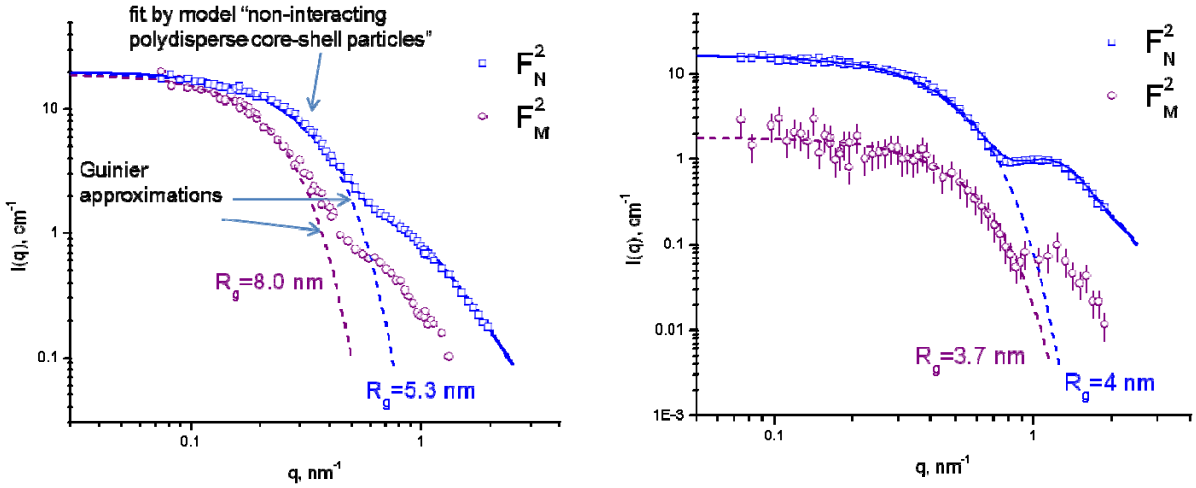
4.3 Polarized neutrons. Magnetic scattering

When polarised neutrons are used with the neutron spins aligned antiparallel (+) or parallel (−) to the direction of an applied magnetic field H scattering intensity depends on the neutron spin state [27,28,19]:

$$I^{+}(q) = F_N^2(q) + [F_M^2(q) - 2F_N(q)F_M(q)]\sin^2\varphi, \qquad (14a)$$

$$I^{-}(q) = F_N^2(q) + [F_M^2(q) + 2F_N(q)F_M(q)]\sin^2\varphi, \qquad (14b)$$

where φ is azimuthal angle in detector plane.



(a) (*b*)

Fig. 12: *Separation of nuclear and magnetic scattering contributions for magnetite ferrofluids based of deuterated cyclohexane and stabilized with oleic (a) and myristic acid (b).*

Thus, character of the scattering is anisotropic now. By making cuts of scattering pattern for both intensities in sectors 0° with 180° and 90° with 270° and solving the system of equations (14) one can separate nuclear F_N and magnetic F_M contributions [29,30]. Separation of nuclear and magnetic scattering on the example of organic ferrofluid is represented in Fig. 12. Different magnetic and nuclear radii calculated from Guinier approximation point to a complex correlation between the magnetic moments of the particles. These correlations differ qualitatively for the two samples [30].
When a polarization analysis of the scattered beam is applied four types of scattering processes are distinguished, two for conserving the neutron spin $(F^{++})^2$ and $(F^{--})^2$ (non spin flip scattering, NSF) and two with reversal of the spin with the scattering $(F^{+-})^2$ and $(F^{-+})^2$ (spin-flip scattering, SF) [27,28,19]. In a particle of radius 4 nm, about 50 % of atoms lie on the surface. The concept of a well-defined super-moment breaks down. Single domain nanoparticles still have uniform magnetization direction, but their spin configuration is not so transparent [31]. In real particles, the surface region thickness is very sensitive to particle shape distortion, surface roughness, surface impurities, defects (like vacancies), compositional inhomogeneity, surface chemical bonds with environment, etc [12]. In literature one can find several possible descriptions of particle spin configuration (see for example [32]).

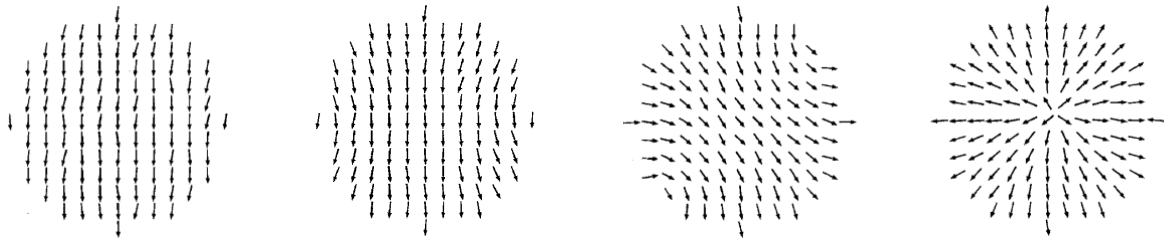


Fig. 13: *Monte Carlo simulation of the possible spin configurations. From [32].*

For the first time the magnetically distinctive core-shell morphology was “seen” in the work [33], where the polarization analysis was successfully applied to directly extract the average spatial distributions of magnetic moments and their correlations. From the obtained scattering in perpendicular direction to the applied magnetic field the authors conclude that magnetic shells of nearly uniform magnitude and direction around the magnetic core exist.

5 Conclusions

Magnetic nanoparticles in general and magnetic fluids particularly nowadays are widely used in technics and are very promising for biomedical applications. For the latter case structure and properties of the substances have to be known beforehand in order to control the synthesis process. The small-angle neutron scattering together with contrast variation technique plays a crucial role in particle structure definition. The newly developed approach of basic functions in contrast variation experiments for the case of polydisperse magnetic nanoparticles opens new possibilities for structural determination. Compared to the monodisperse case, the contrast dependence of scattering integral parameters for the polydisperse system contains a number of additional parameters comprising information about polydispersity function of the

system; their analysis increases the reliability of experimental results. As to the SANS data fitting according to supposed particle models the contrast variation technique results in additional parameters, which can be included in the fit function as constants, thus reducing the number of free parameters and increasing the reliability of the fit.

Acknowledgements: the author would like to thank colleagues without whom this work would have not been possible: M.V. Avdeev, V.L. Aksenov (FLNP JINR, Russia), L.A. Bulavin, V.I. Petrenko (Physics Department KNU, Ukraine), L. Vekas, D. Bica [1952-2008], O. Marinica (Laboratory of Magnetic Fluids CFATR, Romania), V.M. Garamus, R. Willumeit (HZG, Germany), E. Dubois, R. Perzynski (UPMC, France).

References

[1] R.E. Rosensweig, Ferrohydrodynamics, Cambridge University Press 1985.
[2] D. Bica, L. Vekas, M.V. Avdeev, et al.: *J Mag Mag Mater* 311 (2007) 17-21.
[3] M.V. Avdeev, V.L. Aksenov, M. Balasoiu, et al., *J Coll Interface Sci* 295 (2006) 100-107.
[4] S. Odenbach (Ed.), Colloidal Magnetic Fluids: Basics, Development and Application of Ferrofluids, *Lect. Notes Phys.* 763, Springer, Berlin Heidelberg 2009.
[5] C. Scherer, A.M. Figueiredo Neto, *Braz J Phys*, 35(3A) (2005) 718-727.
[6] N.A. Brusnetsov, T.N. Brusnetsova, E.Yu.Filinova, et al., *J Mag Mag Mater* 311 (2007) 176-180.
[7] C. Sun, J.S.H. Lee, M. Zhang, *Advanced Drug Delivery Reviews* 60(11) (2008) 1252-1265.
[8] Q.A. Pankhurst, J. Connolly, S.K. Jones and J. Dobson, *J Phys D Appl Phys* 36 (2003) R167-181.
[9] N.A. Brusentsov, L.V. Nikitin, T.N. Brusentsova, et al., *J Mag Mag Mater* 252 (2002) 378-380.
[10] T.N. Brusentsova, V.D. Kuznetsov, *J Mag Mag Mater* 311 (2007) 22-25.
[11] D. Sellmyer, R. Skomski (Eds.), Advanced Magnetic Nanostructures, Springer 2006.
[12] S.P. Gubin (Ed.), Magnetic Nanoparticle, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 2009.
[13] C.A. Ross, *Annu Rev Mater Res* 31 (2001) 203-235.
[14] J.L. Dormann, D. Fiorani, E. Tronc, Magnetic Relaxation in Fine-Particle Systems, in I. Prigogine, S.A. Rice (Eds.), *Adv Chem Phys* 98, John Wiley & Sons, Inc., Hoboken, NJ, USA 2007.
[15] G. Porod, General Theory, in O. Glatter, O. Kratky (Eds.), Small Angle X-ray Scattering, Academic Press Inc. (London) Ltd., 1982.
[16] L.A. Feiging, D.I. Svergun, Structure Analysis by Small-Angle X-Ray and Neutron Scattering, Plenum Press, New York 1987.
[17] H.B. Stuhrmann, Contrast Variation, in O. Glatter, O. Kratky (Eds.), Small Angle X-ray Scattering, Academic Press Inc. (London) Ltd., 1982.
[18] M.V. Avdeev, *J Appl Cryst* 40 (2007) 56-70.
[19] A. Wiedenmann, *Physica B* 356 (2005) 246-253.
[20] A.V. Feoktystov, M.V. Avdeev, V.L. Aksenov, et al., *Solid State Phenom* 152-153 (2009)

186-189.
[21] A.V. Feoktystov, L.A. Bulavin, M.V. Avdeev, et al., *Ukr J Phys* 54 (2009) 266-273.
[22] M.V. Avdeev, B. Mucha, K. Lamszus, et al., *Langmuir* 26(11) (2010) 8503-8509.
[23] M.V. Avdeev, V.L. Aksenov, M. Balasoiu, et al., *J Coll Interface Sci* 295 (2006) 100-107.
[24] J. Skov Pedersen, *Adv Coll Interface Sci* 70 (1997) 171-210.
[25] R. Massart, E. Dubois, V. Cabuil and E. Hasmonay, *J Magn Magn Mater* 149 (1995) 1-5.
[26] M.V. Avdeev, E. Dubois, G. Mériguet, et al., *J Appl Cryst* 42 (2009) 1009-1019.
[27] R.M. Moon, T. Riste, W.C. Koehler, *Phys Rev* 181(2) (1969) 920-931.
[28] R. Pynn, J.B. Hayter, S.W. Charles, *Phys Rev Lett* 51 (1983) 710-713.
[29] M.V. Avdeev, M. Balasoiu, V.L. Aksenov, et al., *J Magn Magn Mater* 270 (2004) 371-379.
[30] M. Balasoiu, M.V. Avdeev, A.I. Kuklin, et al., *Rom Rep Phys* 58(3) (2006) 305-311.
[31] R.A. Kodama, A.E. Berkowitz, *Phys Rev B* 59 (1999) 6321-6336.
[32] Y. Labaye, O. Crisan, L. Berger, et al., *J Appl Phys* 91 (2002) 8715-8717.
[33] K.L. Krycka, R.A. Booth, C.R. Hogg, et al., *Phys Rev Lett* 104 (2010) 207203.

E 7 Energy materials[1]

R. P. Hermann
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 43^{rd} IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Energy today and tomorrow

Achieving sustainable energy economics is arguably the most important challenge mankind will face in the 21st century. These economics are essentially defined by the offer, the demand, and boundary conditions. In the light of the current demographic growth and the industrial development of emerging countries, an increasing demand is certain at least for the next 50 years. The offer is a matter of debate and financial costs play and will play an essential role. The boundary conditions are essentially set by political decisions that reflect growing ecological concerns. In order to identify the problems and the needs for intense research on energy materials, it is important to set the stage and to consider the current global landscape for energetics and the current policies, notably at the level of the European Union, that aim to define how this landscape *should* look on the horizon of 2050. In this chapter, as the breadth of this topic is daunting and must be limited, after discussing the global aspects that help to identify the important problems, only a few choice studies that have involved x-ray and neutron techniques will be presented.

When consider global energetic problems, the proper scale must be set. A simple estimate of the scale for energy demand is obtained by considering that on the order of 10 billion people live on earth, and that these people have energy requirements of between 1000 and 2000 kcal per day. This corresponds to a power of 50-100 W/person, or 0.5-1 TW worldwide. Regardless of any other fancy activities such as transportation, entertainments, etc... the TW is thus the scale of things when discussing global energy issues. Note also that it takes about 10-20 times more energy to produce food (from fertilization and growth all the way to distribution and storage) than the energy recovered by eating that food[1] in industrialized countries. And indeed, considering the 2009 worldwide primary energy supply of 12150 Mtoe (million tons oil equivalent) and the final consumption of 8350 Mtoe[2], energy rates of 16 and 11 TW are obtained, respectively (1 Mtoe/hour = 11.63 TW), see Fig. 1. By 2050, the industrialization and the population growth will lead to a projected primary power consumption of $\sim$28 TW [1, 3]. This is thus the scale of the demand.

The offer for energy comprises several sources: fossil fuels, nuclear fuel, biomass, hydroelectric, and various other renewables. The first but by far not most serious worry that we might have, would be whether the demand can be met. Indeed, history is full of examples revealing that resource shortage is a strong driver for unrest and war. There is however no real need for worry here. If we consider reserves that are proven, with confidence enough to trade on the stock market, there is globally about 40-80 years of oil, 60-180 years of natural gas, and more than 200 years of coal to meet the demand. The estimated resources, based on geological surveys, are larger and amount to 50-150 years of oil, 200-600 years of natural gas, and more than 2000 years of coal[1]. On these grounds, there is thus no accurate worry to have. The second problem to consider, is that with the current energy mix and the increasing energy needs, the CO_2 concentration in the atmosphere would reach 750 ppm by 2050[3]. From historical records, this concentration is known to have been between 200 and 300 ppm in the last 600000 years[1]. Within this scenario, the CO_2 concentration will by twice larger than in the past million year and likely larger than in the past 100 million years. To put it blandly, in the words of N. Lewis, we are conducting “an experiment, the biggest experiment that humans will ever have done”. The detailed long term consequences of such increasing greenhouse gases on the biospheres are somewhat unclear, but most models predict unpleasant consequences and effects lasting at least for a similar time as the CO_2 mixing time withing the atmosphere/ocean system, of about 1000 years.

Because the raw energy production costs are still strongly biased in favor fossil fuels, with prices

of around 0.02-0.05 €/GJ of heat, or 0.01-0.1 €/kWh of electricity, very comparable to nuclear power plant produced electricity, it is clear that without incentives or boundary conditions, the market is not bound to evolve towards renewable energies. This is best seen by considering that in China the $\sim$10%[4] growth in primary energy consumption is largely matched by construction of, truly said highly efficient, coal fired power plants at an approximative rate of 1-2 plants of 500 MW per week[5]. India, with a similar population faces the same challenge. It is thus certainly not only a matter of scientific research but also a matter of policy that will determine the future developments.

In this respect, the European Union has set down a Roadmap for 2050, that targets a rather aggressive goal: reduce greenhouse gas emissions to 80-95% below the levels of 1990, while improving competitiveness and security of supply[6]. It is however questionable whether such goals would be realistic and even nearly achievable for developing country. The roadmap leaves each country responsible for the detailed choices, but sets goals and describes a set of scenarios based on different shares of renewables, energy efficiency, and new technologies. In all decarbonisation routes, a set of common points define the key outcomes: increased demand for renewable energy, importance of energy savings, increased role for electricity, increased capital investments, decrease in total fossil fuel costs. The five decarbonisation scenarios to be compared with the current trends are a) high efficiency, with $>$40% decrease in demand due to very high savings, b) new technology competition, market driven, without specific support, c) high renewable resources covering 75% of the total and 97% of the electricity consumption, d) carbon capture and storage with up to 18% of nuclear energy in primary consumption, e) low nuclear energy mix with higher carbon capture and storage.

Note that the EU commission is neutral on the use of nuclear power, but that specific countries, such as Germany, have decided upon an aggressive low nuclear scenario following the Fukushima incident in March 2011. In January 2011, 25% of the electricity in Germany was produced from nuclear energy, but nuclear energy amounted only to 10% of the gross consumption. Depending on the specific scenario, renewables should cover 55-65% of the energy by 2050 and enhanced efficiency should yield a reduction in energy consumption between 30 and 40 %. Interestingly, in Europe, the projected purely economic costs are roughly balanced between a 'business as usual' scenario and decarbonisation, because the heavy upfront investments costs are eventually offsetting the raising fossil fuel costs.

2 Technological requirements and fundamental limitations

In order to understand the important challenges and the major components in a -hopefully- sustainable future energy landscape, it is informative to consider the current energy sources and uses. The energy sources for selected countries in Europe and the consumption per sector of activity are presented in Fig. 2. At the European scale, roughly one quarter of the energy is dedicated to industry, one third to transportation, one quarter to domestic use in households. For the sector of transportation, increasing efficiency is well underway, and new technologies are emerging, maybe with the exception of lacking viable alternative concepts for trucks. Interestingly, it is the sector of domestic use that is likely to prove most problematic, as this is fundamentally a distributed problem with a large time constant. In particular in cities much is to be done, as about 2/3 of the energy from coal plants are lost as waste heat. As stated by P. Lyman, CEO Siemens Ltd. "It is increasingly clear that the battle for environmental sustainability will be won or lost in cities. Over half of the world's population now live in urban areas,

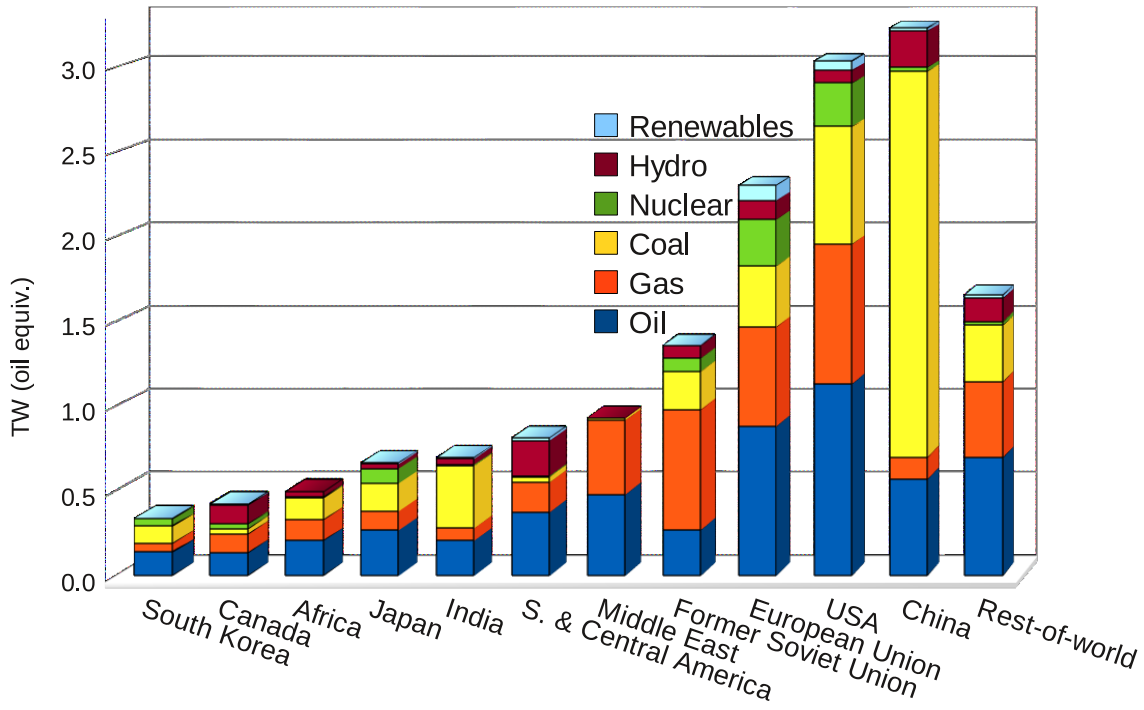


Fig. 1: *Total primary energy consumption in 2010 rate by fuel type* ($1,33.10^{-3}$ *TW assumed equivalent to 1 million tons of oil/year). Source: Ref. [4].*

a figure which will reach almost 60 percent by 2025. Cities consume 75 percent of the world's energy and produce 80 percent of all greenhouse gas emissions. With increasing urbanisation, every city faces a major challenge to achieve its sustainability targets."[7]

Based on the needs and planned development, the European Union has thus defined Key priorities: a) the efficient use of resources at all levels, including energy management in buildings, b) the development and integration of renewables and biofuels in the distribution grids, c) an improved management of waste, in particular of carbon, d) electrification of transportation, e) development of next generation of fission and fusion power plants.

Fundamentally, energy problems are extremely often materials problems, with the notable exception of smart energy grid management, which is a major challenge for information technologies and system engineering, as the integration of intermittent renewable energies will put strain on the energy transportation grid.

Before considering the particular material science problems at which neutrons and x-ray techniques are good at, it is informative to briefly list and discuss what energy sources and technologies there are at our disposal, and how much they could ultimately contribute to a global energy economy, if all the nitty gritty technological details were solved. In the analysis below, I strongly rely on Nathan Lewis' article "Powering the planet" and strongly recommend it to the interested reader[1].

The first energy "source" to consider are the fossil fuels, although to be precise, those fuels as well as hydrogen are better described as energy vectors, as they allow for the transportation and (long term) storage of primary solar energy. These fuels are well known and used, and as discussed above, with the right pricing, there is no real risk of medium term shortage, as the worldwide coal reserves are considerable. However, this source is certainly not a sustainable so-

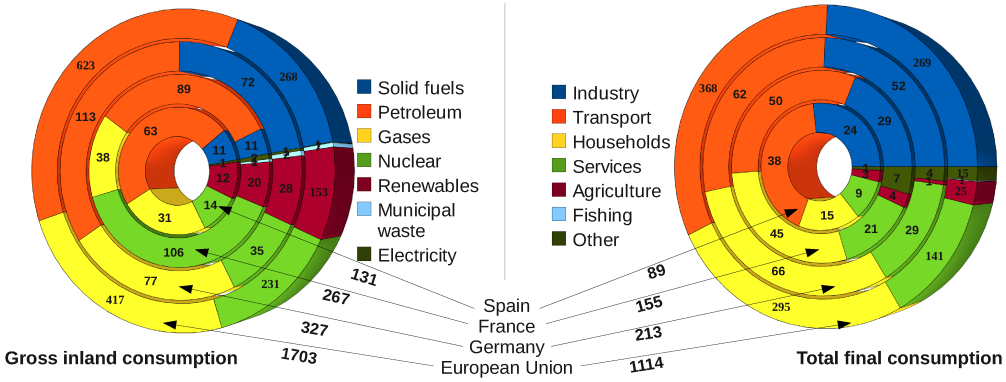


Fig. 2: *Gross inland consumption by energy source for the European Union as a whole, and as example for Germany, France and Spain, left. Total final consumption per sector, right. All numbers in Mtoe for the year 2009. Source:* `www.energy.eu/country_overview`.

lution due to the greenhouse gas emissions. The same is true for methane hydrate clathrates[8], $i.e.$ methane molecules encaged in a crystalline structure by water molecules, with global reserves estimated between one and five millions km^3, which could amount to one half of the global fossil fuel reserves[9]. Using fossil fuels is still a possible approach in conjunction with carbon sequestration technology. The fuel could be converted to non-carbonated fuel, electricity or hydrogen, at central locations and the CO_2 concentrated and buried or converted to carbonates. The option should be explored, but several difficulties are to overcome: the leak rate from geological burial sites, the global pH change in case of ocean burial, and the complete overhaul of the energy distribution system.

The second source would be nuclear power. If we were to suppose that all safety issues are solved, would this be a viable source of energy for mankind? A modern typical nuclear power plant produces energy at a 1 GW rate. If we were to cover the global demand $only$ be nuclear power, by 2050 approximately 20000 plants would be required, that is we need to construct one every second day. The operation time of the plants being about 50 years, we would then start to replace them. It is not technologically impossible, but would consume in about 50 years the total continental uranium reserves, and we then would need to extract uranium from seawater, which would provide about 600 years of power. Or we could search for alternative cycles based on thorium or plutonium. The advantage is that nuclear power does not contribute (directly) to greenhouse gas emissions; as such it is certain that it will play a role in the future, at least as an interim solution, but it is likely not a long term option.

Next, let's turn to the sustainable sources, that is the carbon free renewables: hydroelectric, geothermal, wind, biomass/photosynthesis, and other solar energy: photovoltaics, solar thermal,....

From a total maximum of 4.6 TW potential energy from hydrological flow (all rivers, but excluding ocean currents), the technically feasible hydroelectric power is around 1.5 TW. Interestingly, in 1997 the installed power was 0.6 TW and produced 0.3 TW in average. In 2010, the average production reached 0.4 TW[4]. As this is relatively mature and economically sound technology, it is not surprising that a large fraction is already harnessed, but it will nowhere suffice to meet increased global demand. In a similar way, ocean energy could be harnessed, that is, one could generate electricity from the waves, tides, currents, from salinity gradient, and from thermal gradient of the sea[10]. The total energy in ocean currents however falls

short, as it has been estimated to 0.5-5 TW[10, 11]; it might be locally economical, but the long term consequences of extracting this energy are largely unknown. Tidal energy has more potential, globally up to 10 TW, but it is rather expensive, and there might be serious ecological consequences for a large scale application.
Wind energy is a rapidly developing technology. The total practically usable continental energy amounts to 2 TW, and up to 10-20 TW might be possible offshore. It is thus an important technology, but deployment on a large scale will involve serious issues for electric transportation and might have an impact on local weather. At the TW scale, a production of 1 million turbines is required, and upscaling the problem is far from trivial as the wind energy is depleted when traveling through large windmill fields.
Geothermal energy is an interesting renewable source, locally, and covers for example 50% of the total primary energy in Iceland in 2004. Globally, however, in order to obtain a sufficient temperature difference, 10 km deep wells need to be dug in rock, and cold water injected, that is converted to steam which runs a turbine; however this cools the rock. Ultimately, the total flux of geothermal energy over land at the surface is $\sim$57 mW/m^2, corresponding to at best 11.6 TW of sustainable global energy.
Biomass energy production is increasingly popular but fundamentally inefficient, about 1% of the received solar power, as the photosynthetic process in plants is optimized for sustaining the plants life and not for transforming and storing energy. Depending on the location and production methods, energy farms might barely break even in industrial countries if considering the total energetic costs, including fertilizers, in growing the plants. Globally, photosynthesis provides 90 TW at the basis of the biosphere. About 30 % of the land needs to be covered by biomass producing farms to achieve the 20 TW scale. Biomass could be a contributor but would not make a serious dent in the bill.
Finally, let us consider solar energy. Out of a gross power of 176000 TW, 120000 TW reach the surface of the earth, and considering favorable sites, estimates average to 600 TW of onshore solar power potential. It is worthwhile to remember that ultimately, except for tidal energy (earth-moon system), nuclear energy (remains of a supernova) and geothermal energy (gravitation and radioactivity) all other energy sources can be traced back to solar energy. In order to produce the required 20 TW, less than 0.2 % of the earth surface is needed. This is equivalent to six times 400 x 400 km^2 if one were to centralize the production. It is still a large number, equivalent in the USA to the surface of all roads, and 10 times the surface of all roofs. A practical example for these latitudes (50 $^\circ$ N): we just installed a roof top photovoltaic installation of 5 kW peak with an estimated annual production of 4000 kWh. The lifetime of such panels is about 20 years, and the end consumer costs about 20000 €, that is 0.25 €/kWh, which is still fine for electricity, but horrendously expensive for thermal energy. As of now, the most efficient solar energy converters involve high temperature concentrators, but this technology is limited to some sites. The major obstacle nowadays for large scale deployment are the costs, at $\sim$5-10 times more per kWh than from fossil fuel, the required transportation and storage (in case you want electricity at night or heat in the winter), and geostrategical consideration (safety of supply). It is then a matter of costs and available technology to choose how to harness the solar power: either photosynthetic biomass production (inefficient), or direct photovoltaic conversion to electricity (expensive), or direct conversion to chemical energy $e.g.$ by water splitting, artificial photosynthesis, methanol production, ... (no mature technology, requires new chemistry).
In summary, powering the 21st century is a multivariable and complex problem and there are many potential energy sources that might be exploited. Unfortunately, fusion reactors are still only a long term perspective, so that for now we are stuck with using everything we can get and

diversify the renewable sources while using fossil and nuclear fuel for short or medium term interim source. The bottom line is that in order to meet the future demand and avoid the worst greenhouse gas scenarios, all improvements in efficiency and renewable technologies must be explored.

3 The many facets of energy materials

When considering the aspects of a global energy economy, the concept of an energy relevant material spans an overwhelmingly large area which ranges from concrete and fertilizers all the way to nuclear fuel and membranes for fuel-cells. In a quite loose classification, one could distinguish structural materials, materials for energy transportation, and what I will call here active energy materials. The latter are the main topic of this chapter.

One could at first glance wonder why structural materials are relevant for energetic consideration. The depth of the energy problem however reveals itself when considering that typical energy requirements to produce concrete amount to $\sim$1 MJ/kg. The global cement production is around 2.10^9 tons per year, which amounts to $>$60 GW. Further, cement production is responsible for 5% of the global CO_2 emitted by mankind[12]. For steel, a total energy of about 25 MJ/kg is required in the production, and 1400 million tons are produced annually, which extrapolates to about 1 TW of power[13]. In both cases, there is significant R&D going on to reduce the required energy consumption and approach the theoretical low limits of around 8 MJ/kg[13]. Structural materials are also playing an important role in reducing waste heat, as they provide effective insulation. In particular light weight polymer insulation materials are getting increasingly important. Understanding the detailed properties of structural materials under extreme conditions is another important aspect, as special tungsten claddings will be required and are developed to withstand the intensive neutron and x-ray radiation flux during operation of fusion power plants[14]. Finally, light weight and high strength materials are of particular importance both for the transportation sector, for example for lighter planes, in jet turbines and wheels for high speed trains, and for turbine blades in power plant technologies . Scattering methods have had and still have some relevance in this process, as they allow for obtaining detailed crystallographic and microstructural data. Furthermore, neutron and high-energy x-ray scattering and imaging owing to their unique high penetration power can, for example, help in assessing the aging and defect formation in metals by stress-strain analysis and by tomography[15].

How can energy be transported? Chemical fuels are transported mechanically, in pipelines or supertankers, and this will likely not change even if renewable and sustainable fuels are used. Electricity transportation involves losses that currently average to about 7%. However, other substantial losses are suffered because for efficient transportation, electricity must be raised to high voltage/low current. Thus, a seemingly somewhat remote topic is highly relevant for energy namely the development of high-T_c superconductors. Room temperature superconductors are still very far off, but since 2008, power distribution through a nitrogen cooled high-T_c superconducting grid is reality in some densely populated areas and an expanding market[16]. The crucial role of scattering methods, in particular neutron scattering, in superconductors research is addressed in another lecture.

We will now consider selected examples covering application of x-ray and neutron techniques to active energy materials used for energy sources and conversion, storage and distribution, improvement of efficiency, and waste management, which are the essentials of the energy economics.

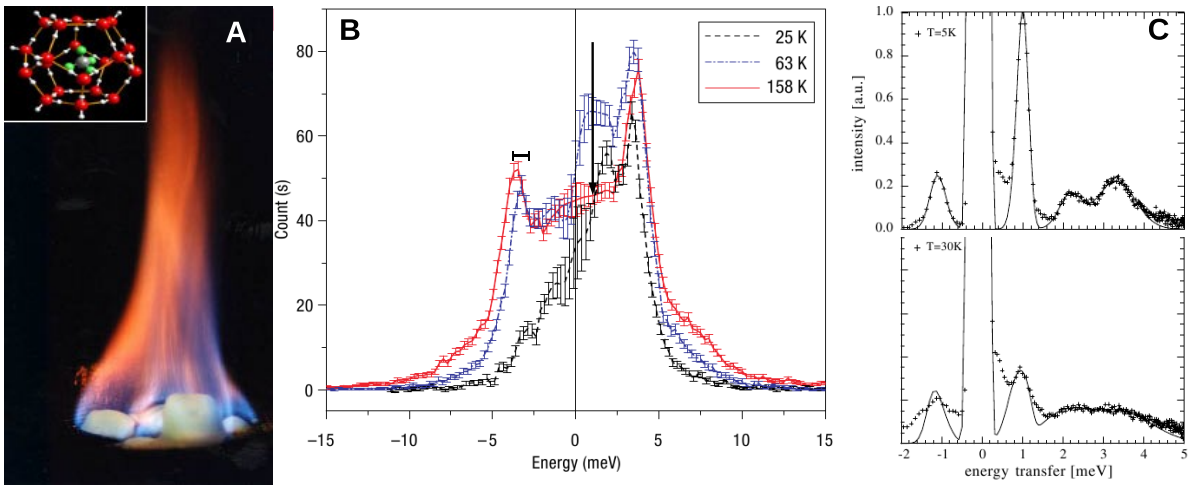


Fig. 3: *Left: burning methane clathrate, inset: structural detail of a CH_4 (in green) enclosed in a $5^{12}6^2$ cage of water molecules (oxygens in red are linked by hydrogen in white). Source: Wikipedia. Middle: ^{83}Kr inelastic nuclear resonance of Kr clathrate hydrate. Note the strong hardening of acoustic phonons below 2 meV and the hardening of the 4 meV mode upon heating [19]. Right: inelastic neutron scattering by natural methane clathrates from the Pacific. At 5 K three distinct inelastic lines belonging to quantum rotational state of methane are observed. The lines broaden with increasing temperature as classical rotational diffusion sets in [20].*

4 Fossil fuels and methane clathrates

Neutron scattering in particular has been extensively used in research of the physics and chemistry of fuels, because of the neutrons unique sensitivity to light elements, in particular hydrogen, the accessibility of large structures, >1 nm, via small angle scattering, and the unique possibilities to match the typical vibrational frequencies in condensed matter.

A typical successful application has been the use of small-angle neutron scattering in order to understand the formation of wax crystals that block the fuel filters in diesel engine[17]. In order to cope with this problem, modern fuels contain a block-copolymer additive that significantly reduces the size of the wax crystals and allows engine operation at low temperatures. The results indicate the formation of supramolecular structures that self-assemble and that the additive allows control crystallization in diesel fuels. The reader is referred to Ref. [18] for a recent review of similar applications.

Because clathrate hydrates, in particular methane clathrates, both constitute a large potential energy source and a tremendous risk in case they destabilize and release methane, their formation, stability, and thermal conductivity has been intensively investigated. They are further involved in clogging of gas pipelines if the pressure and temperature conditions are right and water vapor is mixed with the gas. Clathrates are crystalline guest-host structures in which guest atoms or molecules are enclosed in a framework of chemically different and more strongly bonded host atoms. The inorganic clathrates with Si or Ge based framework and alkaline or alkali earth guests are among the simplest clathrates and have interesting thermoelectric properties. Hydrate clathrates in which water forms the framework and various noble gases or molecules such as methane, hydrogen, or carbon dioxide are guests, are also prominent. Methane clathrate can be harvested from the sea floor on the continental shelf, but rapidly decompose at ambient pressure and temperature. Although mostly comprised of water, they can be ignited by a flame and the

released methane burns, see Fig. 3A. The thermal conductivity of methane clathrates is quite low, between 0.25 and 0.68 W/m/K, much lower than that of ice, about 2.2 W/m/K[8]. Although there is no strong intrinsic disorder in the structure, the clathrate crystals exhibit a low thermal conductivity more akin to a glassy material. The low thermal conductivity is not fully quantitatively understood, but has been linked with the coupling of the guest vibrational modes with the hydrate lattice. Inelastic scattering technique, such as inelastic neutron scattering, inelastic x-ray scattering and inelastic nuclear resonance scattering[19] have been instrumental in identifying the vibrational modes and their anharmonicity, see Fig. 3B. Inelastic neutron scattering has also been instrumental in identifying the rotational modes of the methane guests and their coupling to the framework[20], see Fig. 3C. The vibrational dynamics of clathrate hydrates is extensively reviewed in Ref. [8].

5 Catalysis

The development of new catalytic material and the understanding of existing natural catalytic processes have a major role to play in the future global energy landscape. Catalysis is nature's way of providing a shortcut for a particular reaction, and in particular it provides alternative reaction mechanisms that lower the activation barrier for a specific reaction, hence increasing the reaction rate. For example, getting a handle on the energy required to have water to react with CO_2 to yield methanol and oxygen would be a game-winning progress. Interestingly nature has found ways to do exactly these kind of processes, and advanced scattering technique now yield insight in the mechanisms.

An important example is the Haber process that transforms nitrogen to ammonia (N_2 + $3H_2$ $\rightarrow$ $2NH_3$). In nature this process relies on the nitrogenase enzyme, and an equally important process, the reversible oxidation of hydrogen to two protons and two electrons relies on hydrogenase, see Fig. 4B and C. Biological nitrogen fixation and the Haber-Bosch process are at the base of the proteins available in human food. The Haber-Bosch process uses natural gas in order to produce ammonia in fertilizer factories, an activity that requires $\sim 2\%$ of the global world energy consumption. The process is particularly energy consuming, as the hydrogen must be first obtained from natural gas. It is also particularly frustrating that bacteria are able to catalyze this reaction with little energy using iron, whereas industrial catalysts require platinum and significantly more energy. Understanding how the nitrogen reduction to ammonia is done by nature at low temperature and pressure is thus a crucial endeavor. This reaction mechanisms has been studied in detail using vibrational spectroscopy with THz radiation and by nuclear resonant inelastic x-ray scattering[21]. Iron nuclear resonance vibrational spectroscopy is particularly useful to probe the active sites in metalloenzymes and is not hampered by the high density of low energy vibrational modes from the more remote parts of the molecule. Identifying vibrational modes such as breathing modes provides useful information for nitrogenase chemistry, because the modes are sensitive to specific binding to the iron centers, see Fig. 4A, and as such give insights into the relation between functionality and dynamics in such complex molecules. The experimental data is crucial in validating computational models. In the same spirit, neutron spin echo spectroscopy has been used to track cleft opening dynamics in alcohol dehydrogenase[22].

The interested reader is referred to Ref. [23] which reviews more generally the applications of neutron scattering in the chemical industry, including catalytic processes.

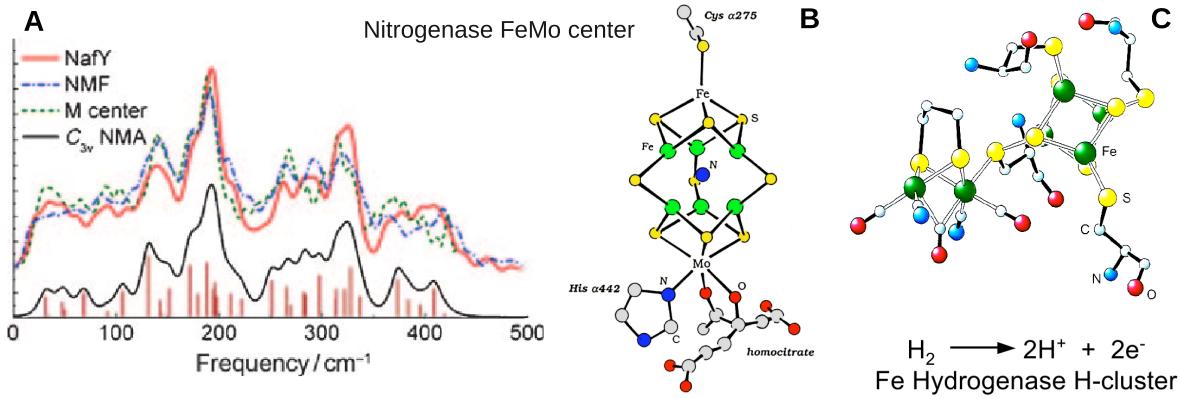


Fig. 4: *A: iron specific vibrational density of states for three forms of the nitrogenase FeMo cofactor and a modeled spectrum from a normal mode analysis for a simplified C_{3v} symmetry (Modified from Ref. [21]). B: nitrogenase FeMo center, C: hydrogenase Fe H-cluster (Source: S.J. George, Lawrence Berkeley National Laboratory).*

6 Solid state direct conversion materials

Thermoelectric materials enable the direct conversion of thermal energy to electric energy by harnessing the heat flow across a thermal gradient to generate a charge flow. Thermoelectric conversion is reversible and relies on the Seebeck, see Fig. 5B, and the Peltier effects for energy harvesting and for thermoelectric refrigeration, respectively. A conversion module is comprise of n- and p-type semiconducting legs coupled thermally in parallel and electrically in series, see Fig. 5A. The efficiency of the semiconducting materials used in such a thermoelectric conversion device is essentially proportional to the product of the electrical resistivity and the squared Seebeck coefficient, and inversely proportional to the thermal conductivity. The latter dependence can be easily understood because the conversion device must maintain a thermal gradient and not create thermal shorts. Currently, for thermal to electric recovery, a realistic device efficiency is between 5 and 10 %, but in theory this efficiency is only bound by the Carnot principle. The field of thermoelectrics has known a regain in interest since the early 1990s for greenhouse gas free and vibration free refrigeration technology. More recently it has been considered as energy recovery technology, on the grounds that about 2/3 of the worldwide energy consumption is lost as waste heat. In that context, thermoelectric conversion, although not extremely efficient, might help in overall energy savings by recovering this waste heat. In particular, in the automotive industry, a 10% fuel economy could be achieved by recovering heat from exhaust gases[24].

Scattering methods have proven extremely useful in understanding the relationship between the structure and the lattice dynamics in thermoelectric materials. There is several options in achieving the required low thermal conductivity while maintaining good electrical conductivity, namely, a) inserting loosely bound guest atoms in cage-like structures, such as Si or Ge based clathrates, b) using materials with very large unit cells, c) using nanostructure materials where grain boundary scattering reduces the thermal transport. A good recent review is given by Snyder *et al.*[25]. Ultimately, thermal transport is governed by the propagation of phonons in the material, and is proportional to the energy they transport, $i.e.$ the materials specific heat, the transport velocity, $i.e.$ the speed of sound, and the mean-free path or average collision time.

PbTe is one of the interesting materials for medium temperature (500-800 K) thermoelectric

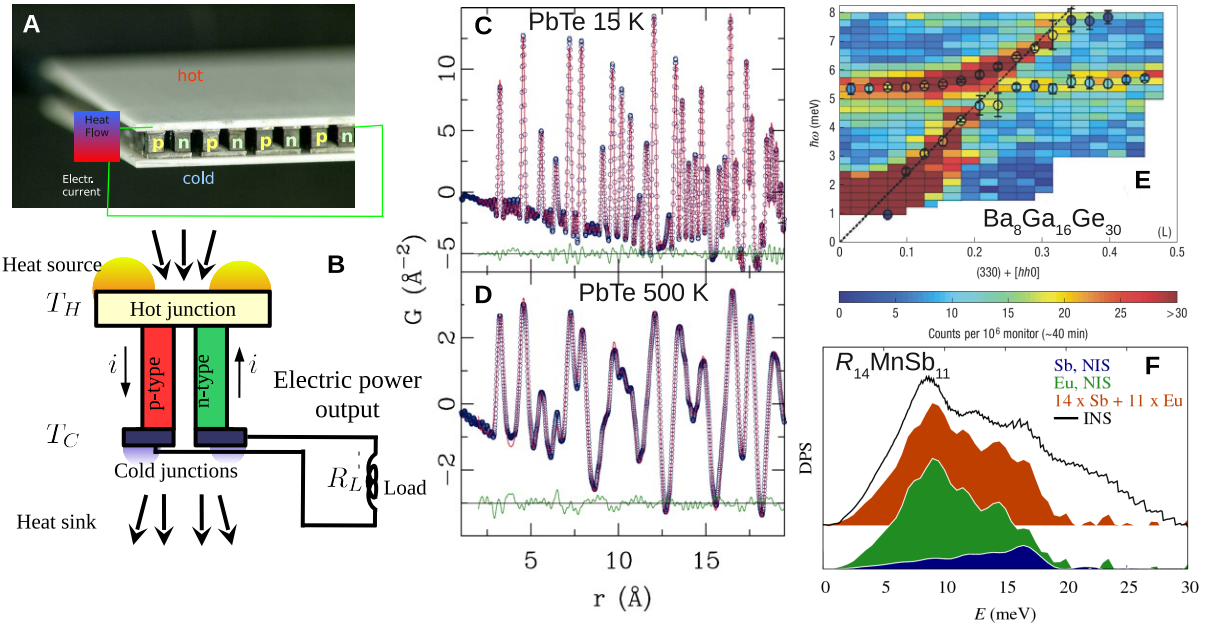


Fig. 5: *A: A thermoelectric module (Source: A. Weidenkaff, EMPA). B: Schematic representation of heat and electric current for the Seebeck effect. C and D: Pair distribution function in PbTe a low and high temperature (Adapted from [26]. E: Phonon dispersion curve in the [ξξ0] direction where a possible avoided crossing is observed[27] in $Ba_8Ga_{16}Ge_{30}$. F: Total density of phonon states as obtained from inelastic neutron scattering compared with the weighted element specific density of phonon states for Sb and Eu in $R_{14}MnSb_{11}$. Eu is used as substitution for Yb as nuclear inelastic scattering is not yet possible for Yb[28].*

applications and as such for medium temperature power generation, such as envisioned by the automotive industry. PbTe crystallizes in the rather simple cubic NaCl structure, and can readily be doped towards p- or n-type by addition of electropositive, Na or K, or electronegative, I or Br, elements, respectively in order to improve its thermoelectric performance. The structure of lead telluride has been recently investigated by total scattering of x-rays and neutrons[26] at low and high temperature, see Fig. 5C and D, and, interestingly, in contrast to conventional phase transitions that lead to distortions and lower symmetry upon cooling, it was found that in PbTe dipole moments appear in the local structure upon heating. This behavior contrasts with the observation of the average structure by diffraction that shows that PbTe adopts the NaCl structure at all temperatures. Because the pair distribution function obtained from total scattering is uniquely sensitive to local distortions it enables to go beyond the average structure. This local distortions are believed to be responsible for the highly anomalous behavior of thermal conductivity and charge carrier scattering. The authors suggest that new thermoelectrics could be found among other materials that, like PbTe, are close to a ferroelectric instability[26].

Inorganic clathrates with Si or Ge framework are also interesting thermoelectric materials, as they both allow for doping and accommodate loosely bound guests in the clathrate void, the vibrational dynamics of which is related to an overall low thermal conductivity between 0.7 and 2 W/m/K at room temperature, depending on the specific structure. The presence of a mostly, but not totally, localized vibrational mode associated with the guest atoms, typically Sr, Ba, or Eu, for Ge clathrates, leads to an avoided crossing between the low lying guest optical phonons and the acoustic cage phonons. This behavior leads to a lowering of the phonon group velocity

and of the thermal conductivity and has been observed by inelastic neutron scattering by large single crystals using a triple axis spectrometer, see Fig. 5E.
$Yb_{14}MnSb_{11}$ is new complex Zintl compound, $i.e.$ a compound that is valence balanced with both ionically bonded atoms and complex polyions of covalently bonded atoms. In $Yb_{14}MnSb_{11}$ the complex polyanionic units $[MnSb_4]^{9-}$ and $[Sb_3]^{7-}$ balance the Yb^{2+} cations and Sb^{3-} anions; the unit cell, containing 208 atoms, is simply huge, with V>6000 Å^3. This complex structure has a very low lattice thermal conductivity <0.4 W/K/m with excellent thermoelectric performance above 1000 K, and was almost immediately adopted by NASA for thermoelectric generators[25]. A combination of inelastic neutron and ^{151}Eu and ^{121}Sb nuclear resonant inelastic scattering, see Fig. 5F, has recently revealed that it is the backfolding of the Brillouin zone associated with such a large unit-cell that leads to an extremely large number of flat optical phonon modes and is responsible for the low thermal conductivity[28].
Next to thermoelectrics, it is worthwhile to mention the magnetocaloric materials that might prove useful for clean and efficient refrigeration technology. The magnetocaloric effect relies on the adiabatic transfer of entropy from the magnetic to lattice (phonon) degrees of freedom. In a refrigeration cycle, a magnetic field is applied to an initially paramagnetic material which orders magnetically. The lost magnetic entropy is transferred to the crystalline lattice which leads to an increase in temperature. The material is then left to cool to ambient temperature and afterwards thermally contacted to the load to refrigerate and the magnetic field is removed. The magnetic entropy thus increases and the lattice entropy decreases, hence causing refrigeration. The material is then decoupled from the load and a new cycle starts. Giant magnetocaloric effects have been obtained for relatively inexpensive Fe and Mn based materials that possess a Curie temperature close to room temperature. Neutron diffraction, with its unique sensitivity to magnetic order has enabled detailed studies of the magnetic and atomic structure above and below the phase transition[29]. Inelastic scattering studies that aim at unraveling the coupling of the magnetic and lattice degrees of freedom are ongoing .

7 Hydrogen storage

The hydrogen economy is sometimes considered as the ultimate solution for green energy. Although this might be true for the end consumer, which uses ambient oxygen and the hydrogen fuel to release energy and water vapor, the global story is more complex. Note that there has been strong criticism against the hydrogen route and that alternative fuels such as methanol might prove more sustainable[30], but as for now, no technology should be left unexplored.
First, the hydrogen will have to be produced, in the ideal cases from water splitting using solar energy, more practically, nowadays, as part of the Haber-Bosch process by using fossil fuels. The thus obtained hydrogen must then be stored for distribution to the end consumer which is both an infrastructure challenge and is at least as dangerous as transporting natural gas. Finally, the end consumer must release the stored hydrogen in a safe manner before generating energy either by relatively inefficient controlled combustion or by using more efficient fuel cells.
As a particular case study we will discuss the controlled release of hydrogen from sodium alanate. A key factor in devising a good hydrogen storage material is the relative hydrogen mass density in the storage material. It is relatively easy to exceed liquid hydrogen storage density, even at ambient temperature, for example in rare earth materials. However, the ratio of contained hydrogen to the container mass is extremely unfavorable. Hence, light materials must be considered, such as ammonia borane (H_3NBH_3) [31] or sodium alanate ($NaAlH_4$)

[32]. Neutron scattering and imaging is the method of choice for visualization of hydrogen, and in particular the large incoherent scattering cross section gives readily access to hydrogen dynamics through measurements of its self-correlation function.

In a recent study, Leon *et al.* have investigated the details of hydrogen release upon heating sodium alanate, a release that occurs in a two step reaction, see Fig. 6 [33]. The authors have used quasielastic neutron scattering at the JCNS backscattering instrument SPHERES in order to follow the decomposition reaction. Specifically, there is four hydrogen bearing species involved in the reaction. In a preliminary measurements, the authors have established that sodium hydride (NaH) has an essentially elastic scattering function, see right of Fig. 6. The scattering for H_2 is essentially not seen and contributes to the background as it diffuses too fast. For the intermediate species Na_3AlH_6 the scattering function exhibits a quasielastic Lorentzian component, see left of Fig. 6. The decomposition reaction of $NaAlH_4$ can thus readily be followed by using Na_3AlH_6 as a proxy and by analyzing the total scattering signal in an elastic Gaussian and a quasielastic Lorentzian part. This study thus gives access to the reaction rates, see middle of Fig. 6, and the authors concluded on the self-catalytic role of Na_3AlH_6 in the hydrogen release mechanism. An extensive review of neutron scattering applied to hydrogen storage materials is made in Ref. [34].

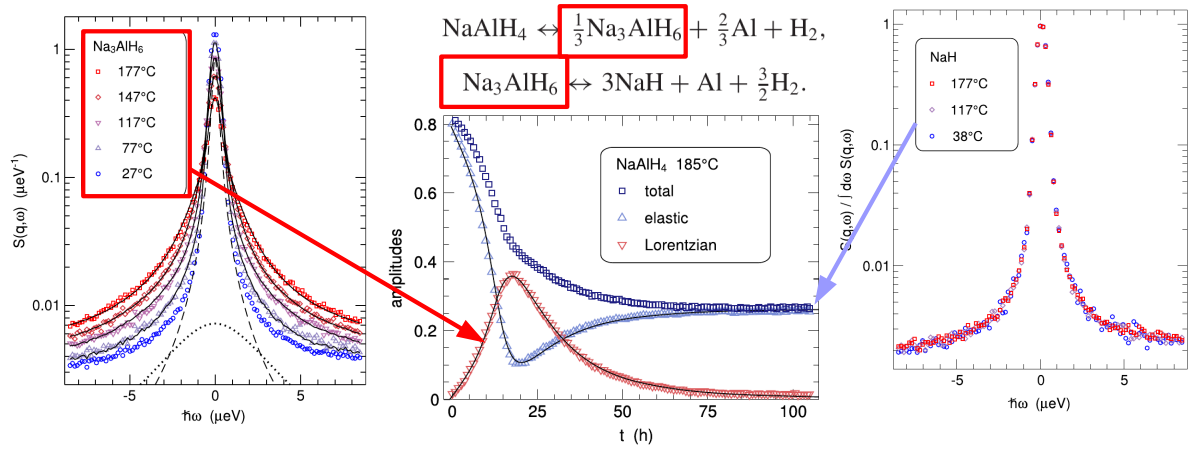


Fig. 6: *Top: two-step reaction for hydrogen release from sodium alanate. Left and right: scattering function for Na_3AlH_6 and NaH, respectively. Middle: time dependent scattering function upon decomposition modeled with the relative abundance of the intermediate and final products (Adapted from [33]).*

8 Photovoltaics

As discussed above, in the long run, harvesting solar energy is simply the way to go as the available energy by far exceeds to sum of all other usable energy sources. There is several ways to more or less efficiently harness solar power, such as biomass production, synthetic photosynthesis or water splitting, and direct photovoltaic conversion to electricity. Electricity is a high-grade form of energy but it only corresponds to about 10 % of the global energetic demand, so that an all-photovoltaic economy would be somewhat over the top and extensive research is certainly required for alternative routes, for example for direct chemical fuel production.

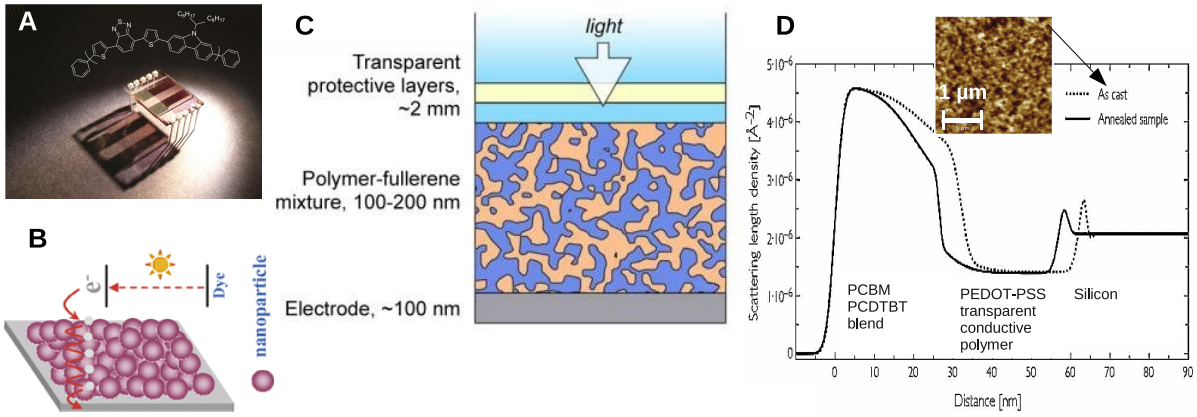


Fig. 7: *A: Demo polymer photovoltaic converter, inset: PCBTBT polymer structure (Source: Science Daily / A. Parnell). B: Dye sensitization by nanoparticle addition[36]. C: Schematic dendritic polymer structure that maximizes surface in order to avoid charge recombination[35]. D: Scattering length density before/after annealing obtained from neutron reflectometry; Inset: SPM image of as cast blend[37].*

Photovoltaics is arguably the most mature of the solar energy technologies at least for large scale applications. The estimated installed power at the end of 2011 reaches 60 GW. The major obstacles for a global use is that it is a quite expensive technology, based on the use of specifically prepared silicon, and that there is no efficient storage mechanism, $i.e.$ the base load must still be covered by conventional power plants. Ultimately one would wish for a very cheap technology, say a photovoltaic paint[1], but on the medium run more pragmatic alternatives to silicon are explored, such as light weight polymer based cells, see Fig. 7A. The core idea for such cells is to sensitize a polymer, for example by including nanoparticular dyes [36], see Fig. 7B, in order to enable efficient charge separation. A major difference between silicon and polymer based cells is the rather low mean free path for charge carriers in the latter. Charge recombination thus must be avoided, that is, the required diffusion path for the charge carriers must be minimized, which can be achieved by a dendritic patterning, see Fig. 7C.

Similarly as for hydrogen storage material, neutron scattering is a powerful technique for investigating hydrogen bearing polymers. In particular, polymer thin films such as used for photovoltaic cells can readily be investigated by neutron reflectometry. Adequate scattering contrast can be achieved by using deuterated or partly deuterated blends. A recent study[37] has shed light on the annealing mechanisms of the polymer blends. In order to improve the efficiency of the P3HT:PCBM benchmark system, thermal annealing is required, an annealing that modifies the vertical structure and improves the efficiency from 2 to about 5%. The authors have used neutron reflectivity and wide angle x-ray scattering under grazing incidence (GIWAXS) in order to explore the formation of nanostructures in a different donor polymer for organic photovoltaic application, PCDTBT, see Fig. 7A inset, blended with a fullerene acceptor PC70BM. The neutron reflectivity study reveals that in the as cast system, the PCBM blend is enriched at the surface and that PCBM migration towards the surface occurs upon mild annealing at 70 °C. The surface roughness itself is not significantly increased, but a slight broadening of the buried interface is observed, see Fig. 7D, that facilitates hole extraction. The GIWAXS measurements reveal that the annealing process does not modify the cristallinity of the PCBM or the size of PCBM crystallites, as no significant peak broadening or sharpening is observed.

9 Fuel cells

Finally, storage and recovery of electrical or chemical energy is an extremely important process. Notably battery technology and lithium ion batteries have been extensively studied, and the interested reader is referred to a review in Ref. [38]. Fuel cell technology is in some sense an alternative to electricity storage in batteries, as in fuel cells it is a chemical energy vector such as hydrogen or possibly methanol, that is oxidized in a controlled way in order to generate electrical current, see Fig. 8A. Note that this fuel cell process is also in principle reversible and could be used *e.g.* for water splitting by injecting electric current. There is many challenging technological aspects that are matter of current research as competitive materials for the anode, the cathode and the electrolytes are needed. In particular, selective permeability to specific gases is required, and management of the generated water is important. Neutron imaging is particularly powerful to visualize water in fuel cells in operando, see Fig. 8B and C.

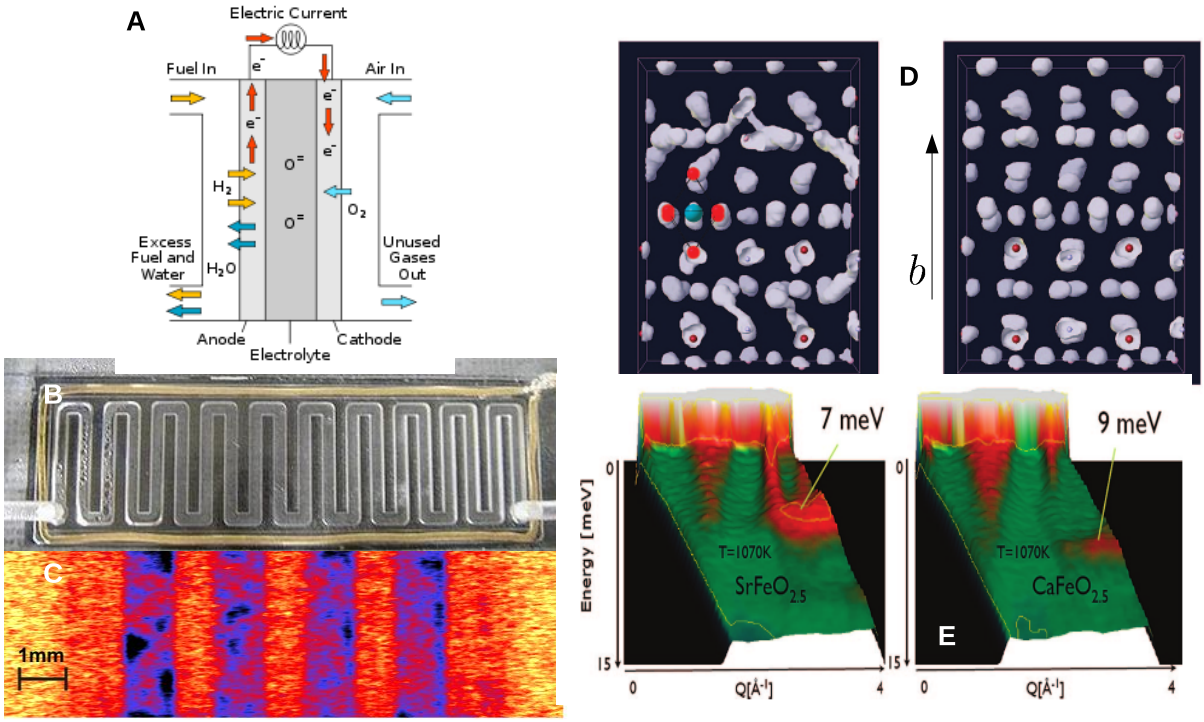


Fig. 8: *A: General principle of a solid oxyde fuel-cell (Source: Wikipedia) B: Water condensation on a proton exchange membrane fuel cell. C: Neutron tomography imaging water in operando in a fuel cell (Source B&C:* `ecl.web.psi.ch/fuelcell/index.html`*). D: Molecular dynamics simulation of oxygen displacements in* $SrFeO_{2.5}$ *(left) and* $CaFeO_{2.5}$ *(right), and E: corresponding inelastic neutrons scattering signal at 1070 K (Adapted from Ref. [39]).*

Ion mobility, and in particular oxygen ion mobility at (relatively) low temperatures (<1000 K) is a key problem for oxygen membranes in solid oxide fuel cells. Typically, oxygen ions move about in solids by a thermally activated process and the ions jump via a vacancy mechanism of diffusion. An inelastic neutron scattering study has recently shed some light on the oxygen diffusion in important iron oxide based materials with nominal perovskite structure ($CaFeO_{3-x}$ and $SrFeO_{3-x}$). The authors have shown[39] how the lattice dynamics trigger the oxygen mobility at low temperatures by combining insights from ab initio molecular dynamics calculations

and inelastic scattering. The limiting factors for oxygen mobility are the vacancy concentration, the barrier height, and the rather large size of the oxygen ions, about 1.4 Å. One key question is why in $SrFeO_{3-x}$ the Brownmillerite structure ($x = 0.5$) can be filled up completely by electrochemical oxidation at room temperature up to the perovskite ($x = 0$) whereas this is possible only under high oxygen pressure for $CaFeO_{3-x}$. The inelastic scattering measurements conducted on IN6, ILL, reveal that in $SrFeO_{2.5}$ a low energy phonon mode is visible at 7 meV for all temperatures, whereas this mode is located at 12 meV in $CaFeO_{2.5}$ and softens to 9 meV only above 1070 K. This soft mode is associated with a structural instability that is triggered above a critical size of the b lattice parameter, a parameter that is always sufficiently large in $SrFeO_{2.5}$ but reaches the critical value only a high temperature in $CaFeO_{2.5}$. Closer inspection of the inelastic scattering data, see Fig. 8E, further reveals that in $SrFeO_{2.5}$ a quasielastic diffusive contribution is seen, that merges with the 7 meV phonon mode, a contribution absent in $CaFeO_{2.5}$. Molecular dynamics calculation helps in understanding this diffusion in revealing that in $CaFeO_{2.5}$ the oxygen atoms vibrate about their well defined position, whereas in $SrFeO_{2.5}$ migration of apical oxygen to vacancies on interplane tetrahedral sites is possible, see Fig. 8D. The authors conclude that the oxygen diffusion mechanisms, that involve a change in Fe coordination, must be attributed to specific lattice vibration modes and are reinforced by the rapid valence fluctuation of iron that acts as a chemical chameleon in changing coordination easily from square-planar (in $SrFeO_2$) through tetrahedral and pyramidal all the way to octahedral (in $SrFeO_3$). Soft phonon mode engineering might thus prove an important tool for more efficient materials.

10 Conclusions

The search for energy, and more particularly, the search for efficient and improved energy materials, are certainly one of the important challenges that mankind will face in the 21st century. In order to avoid critical levels of greenhouse gas in the atmosphere, decarbonisation of the energy is simply a must. Balancing this decarbonisation with economic sustainability is however challenging in particular for developing countries. Fortunately, in the long run at least, there is sustainable energy solutions, most prominently by harnessing solar energy.

Intense research, notably in material science, is required in order to improve upon efficiency in existing technologies and discovered new avenues. Selected examples have been presented here, using diverse techniques such as total scattering pair distribution function, nuclear resonance inelastic scattering, inelastic and quasielastic neutron scattering, etc... but even more techniques that are important have not been presented, such as x-ray absorption fine and near-edge structure spectroscopy, x-ray anomalous diffraction. Advanced scattering and imaging using x-rays and neutrons are bound to play an important role as they ultimately allow us to have a look inside the materials and unravel the interplay between structure, functionality, and dynamics, by revealing where the atoms are and how they move.

References

[1] N. Lewis, MRS Bulletin **32**, 808-820 (2007); N. Lewis, Eng. & Science **2** 12-23, (2007). See also nsl.caltech.edu/energy

[2] International Energy Agency, www.iea.org/stats/index.asp

[3] M. Hoffert *et al.*, Nature **395**, 881-884 (1998).

[4] Statistical review of world energy 2011,www.bp.com/

[5] www.powermag.com/coal/Debunking-the-Chinese-coal-monster-myth_1527.html

[6] www.roadmap2050.eu/

[7] www.siemens.ie/_documents/Siemens_NW_IRE_030210.pdf

[8] M. M. Koza and H. Schober, "Vibrational Dynamics and Guest-Host Coupling in Clathrates Hydrates", in L. Liang *et al.* Eds., Neutron Applications in Earth, Energy and Environmental Sciences, Springer (2009).

[9] A.V. Milkov, Earth-Science Review **66**, 183197 (2004).

[10] www.eu-oea.com/index.asp?bid=232

[11] ocsenergy.anl.gov/documents/docs/OCS_EIS_WhitePaper_Current.pdf

[12] J. Vanderley, UNEP - Industry and Environment **26** 62-63 (2003); J. Struble and J. Godfrey, International Workshop on Sustainable Development and Concrete Technology, Beijing, pp.201-211 (2004).

[13] J. de Beer *et al.*, Annu. Rev. Energy Environ. **23**, 123205 (1998).

[14] www.sciencedaily.com/releases/2010/03/100304102322.htm

[15] E. Lehman, "Imaging and applications", in L. Liang *et al.* Eds., Neutron Applications in Earth, Energy and Environmental Sciences, Springer (2009); M. Grosse and P. Ottlinger, Mat. Sci. and Engin. A **436**, 88-92 (2006).

[16] www.economist.com/node/21540385

[17] http://www.ill.eu/industry/solutions/applications/soft-materials/flow-improvers-for-diesel-fuels/

[18] Y. B. Melnichenko and G. D. Wignall, J. Appl. Phys. **102**, 021101 (2007).

[19] J. S. Tse *et al.*, Nature Materials **4**, 917 (2005).

[20] C. Gutt *et al.*, Europhys. Lett. **48**, 269 (1999).

[21] Delfino E. *et al.*, Angew. Chem. Int. Ed. **49**, 3912-3915 (2010).

[22] R. Biehl *et al.*, Phys. Rev. Lett. **101**, 138102 (2008).

[23] P. Albers and S. Parker, "Applications of neutron scattering in the chemical industry", in L. Liang *et al.* Eds., Neutron Applications in Earth, Energy and Environmental Sciences, Springer (2009).

[24] *Thermoelectrics Handbook, Macro to Nano*, ed. D. M. Rowe (Taylor and Francis, Boca Raton, 2006); R. Hermann, C6 Thermoelectrics, 41st IFF Springschool 2010 Lecture Notes.

[25] G. J. Snyder and E. Toberer, Nature Materials **7**, 105-114 (2008).

[26] E. S. Bozin *et al.*, Science **330**, 1660-1663 (2011).

[27] M. Christensen *et al.*, Nature Materials **7**,811-815 (2008).

[28] A. Moechel *et al.*, Phys. Rev. B **84**, 184303 (2011).

[29] K. A. Gschneidner Jr. *et al.*, Rep. Prog. Phys. **68** 1479-1539 (2005); L. J. Liu *et al.*, Powder Diffr. **25**, S25 (2010).

[30] G. A. Olah *et al.*, Beyond Oil and Gas: The Methanol Economy, Wiley (2006).

[31] F. Stephens *et al.*, Dalton Transactions **40**, 2613-2626 (2007).

[32] C.P. Baldé *et al.*, J. Am. Chem. Soc. **130** 67616765 (2008).

[33] A. Leon and J. Wuttke, J. Phys. Cond. Matter **23**, 254214 (2011).

[34] M. Celli *et al.*, "Hydrogen and hydrogen storage materials", in L. Liang *et al.* Eds., Neutron Applications in Earth, Energy and Environmental Sciences, Springer (2009).

[35] `www.maths.ox.ac.uk/groups/occam/research/resources-energy-and-environment/ree9`

[36] S. H. Kang *et al.*, Adv. Mater. **20**, 54 (2008).

[37] P. Staniec *et al.*, Adv. Energy Materials **1**, 499 (2011).

[38] M. Catti, "Lithium Ion Materials for Energy Applications", in L. Liang *et al.* Eds., Neutron Applications in Earth, Energy and Environmental Sciences, Springer (2009).

[39] W. Paulus *et al.*, J. Am. Chem. Soc. **130**, 16080 (2008).

E 8 Structure and dynamics of proteins

R. Biehl

Jülich Centre for Neutron Science 1 & Institute of Complex Systems 1

Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Proteins are biological macromolecules present in all cells and in body liquids. They work as nanomachines of live to produce material, move objects to the place where they are needed, degrade toxic chemicals, regulate the velocity of processes or protect the cell e.g. from viruses as part of the immune system. Proteins are synthesised as linear polypeptides of 21 different amino acids that are connected by peptide bonds. The primary structure of the protein as the sequence of amino acids is coded in the sequence of DNA. Messenger RNA is an information copy of the DNA piece containing the protein sequence. In the ribosome, a complex of RNA and ribonucleoproteins, proteins are synthesised as defined by the messenger RNA. During the protein synthesis the protein strand folds to a unique 3 dimensional structure. Sometimes helper proteins, the chaperons, are necessary to find the correct folding structure. Misfolded or damaged proteins are degraded in the proteasome to smaller pieces and the amino acids are reused to build new proteins.

During the folding process the secondary structure evolves with α-helices as twisted structures and β-sheets as flat structures or disordered regions between them. α-helices and β-sheets are mainly stabilised by hydrogen bonds between specified neighbours e.g α- helices share a hydrogen bond between the N-H and the C=O of the amino acid 4 amino acids earlier in sequence leading to the helical structure. The spacial arrangement of these structures is named tertiary structure. If oligomers are build of several monomers the structure is called quaternary structure. Which structure evolves is dependent on the properties of the amino acid side chains in sequence. The side chains make the amino acid a weak acid or base or hydrophilic or hydrophobic, if the side chain is polar or nonpolar. Hydrophobic amino acids are mainly found the interior and hydrophilic amino acids are mainly found at the surface of globular proteins with contact to the surrounding water. This stabilizes these proteins against unfolding because the hydrophobic contact to water is minimized. The acidic or basic character of the amino acids at the surface determines the charge of the protein. If the net pK of the protein is equal to the pH of the buffer the protein is neutral. If the protein pK is different from the pH of the buffer the protein is charged and stabilised by coulomb repulsion. Differently charged patches at the protein surface can lead to aggregation or stabilise the tertiary structure of oligomers. Inside of the protein e.g. at the active centre the character of the amino acids is used to degrade chemicals or to protect them from hydrolysis.

A unique source of known protein structures is the Protein Data Bank (PDB) [1]. 90% of the accessible structures are determined by X-ray crystallography, while 9% were determined by NMR spectroscopy. The large number of known structures in the PDB should not emphasis that all proteins are as rigid as suggested from structures in the PDB. Intrinsically Unstructured Proteins (IUP)[2] are a class of proteins that have no fixed tertiary structure but make up to 30% of all proteins. IUP's challenge the structure-function paradigm and are rarely found in the PDB because they don't form a crystal. IUP's have a defined amino acid sequence but no or only partially secondary structure connected by disordered regions. Some IUP's fold to a defined structure, if they are bound to their target. This can be a substrate or another protein, where the unfolded structure allows a better adaption to the target than a larger structured protein [3]. Unfolded parts of a protein with stable tertiary structure, can be

used e.g. for autoinhibition by binding of the unfolded part at the active centre [4], as it is found in the case of Vav Ac domain, where phosphorylation of the unfolded terminal prevents binding to the active site and enables function.
An early model about protein specificity was the “lock and key” model, which assumes an exact fit of the protein active site to the substrate due to complementary geometrical shapes [5] as found in crystal structures. To explain also the stabilisation of the transition state with a bound substrate in different configuration compared to the unbound state the “induced fit” model [6] allows a reshaping of the binding site to the substrate including local configurational changes of amino acids or large structural changes as for allosteric transitions.
Today it is realised that proteins are quite flexible objects, which show dynamics on all length scales. The fastest motions are bond vibrations, side chain rotation at the protein surface or torsion of buried groups with sub-angstrom amplitudes on picosecond timescales. Slower motions on nanosecond timescales with angstrom amplitudes are relative motions of complete domains as hinge bending movements or swapping of domains. Allosteric transitions, folding and unfolding will happen on microsecond timescale and nanometer length scales. Protein function is linked to these internal dynamics. Often the accessibility of the active centre is only possible if domains swap away allowing removal of the product and binding of a new substrate. Rearrangements of amino acids to adjust the orientation of functional groups requires local flexibility of neighbouring amino acids.
In the following we will address two important topics. First the determination of the large scale structure, respectively the shape of proteins is described. Secondly the dynamics from atomic vibrations to large scale domain motions is reviewed and how it can be accessed by scattering techniques.

2 Large scale structure determination

The protein primary structure is determined by aligning fragments of the complete sequence [7]. Edman degradation [8], splits of the terminal amino acid of a fragment for analysis and this step is repeated for the complete fragment. Another method uses Tandem Mass Spectroscopy to analyse the mass of a fragment and to determine in a second step after splitting into amino acids the fragment composition. Secondary structure composition of the complete protein can be determined by Circular Dicroism Spectroscopy, using the different absorption of left and right polarised light of secondary structure elements, or Infrared Spectroscopy, which uses the change of the amide I and amide II bands around 1600cm^{-1}. The IR bands of the peptide N-H and C=O bonds, mainly contributing to these bands, are shifted due to hydrogen bonds, which stabilises the secondary structure [9].
X-ray crystallography for atomic structure determination requires a crystal of about 100µm in size. The proteins are restricted in their motion by the neighbours in the crystal even if the water content of some protein crystals reaches 50%. NMR measures the magnetic resonances of atoms, which depends on their local magnetic environment in the protein, and reconstructs the protein structure. NMR is restricted to smaller proteins because for larger proteins the similarity of local atomic environments becomes more likely and the resonances start to overlap and inhibit reconstruction.
The large scale structure of multidomain proteins and protein complexes without the geometrical restrictions of the crystal structure can be determined by Small Angle Scattering methods with X-ray or neutrons (SAX and SANS).

2.1 Small Angle scattering by X-rays and neutrons

X-rays are scattered at the electron cloud around the nucleus, while neutrons are scattered at the nucleus dependent on isotope and spin orientation. Scattering techniques depict the changes in the contrast, which is the difference in scattering length density between protein and solution. For X-rays the contrast is due to a change of the number of electrons in the electron cloud of individual atoms. For organic systems hydrogen atoms have the largest total neutron scattering length and therefore hydrogen density gives the largest contribution to the neutron scattering contrast. To have a maximal contrast between solvent and protein the measurements are done in D_2O instead of H_2O because deuterium has a much smaller scattering length compared to hydrogen. About half of the protein atoms are hydrogen atoms with the main contribution to the total protein scattering. As a consequence of using D_2O protons bound in -O-H, -S-H groups and a part of the peptide N-H groups exchange the proton with deuterium. About 10% of the hydrogen is exchanged in this way and reduces the local contrast to the solvent D_2O buffer.

The protein shape and local secondary structure arrangements mainly determine the small angle scattering of proteins. Figure 1 shows the expected neutron scattering according to the structure in the PDB for a concentration of 10 mg/ml in a D_2O buffer for 3 different sized proteins. The coherent + incoherent background scattering of e.g. 100 mM NaCl in the buffer has a contribution of about 0.012 cm^{-1} to the D_2O background scattering of about 0.052cm^{-1}. The protein incoherent scattering is mainly due to the non-exchangeable hydrogen. This additional constant incoherent background does not exist for SAX measurements.

At low Q the coherent scattering I(Q=0) is mainly determined by the contrast and the mass of the protein. In the Guinier regime with $Q<1/R_g$ the scattering is described by the radius of gyration R_g according to $I(Q)\sim\exp(-R_g^2Q^2/3)$. Above this regime the scattering is more influenced from the detailed shape of the protein. For ADH (see figure 1) a shoulder above 1 nm^{-1} appears indicating the tetrameric structure of ADH, which is composed of two dimers arranged in a crossed configuration. The position of the shoulder is attributed to the distance of the monomers. In general the scattering of smaller proteins shifts in this region to larger Q as observed for Rnase. Around 2-3 nm^{-1} the coherent scattering equals the incoherent protein scattering. For larger Q the incoherent scattering dominates and the coherent contributes only about 20% to the total signal. Above 0.7 nm^{-1} an additional structure is visible in the coherent scattering, which is due to secondary structure elements and their

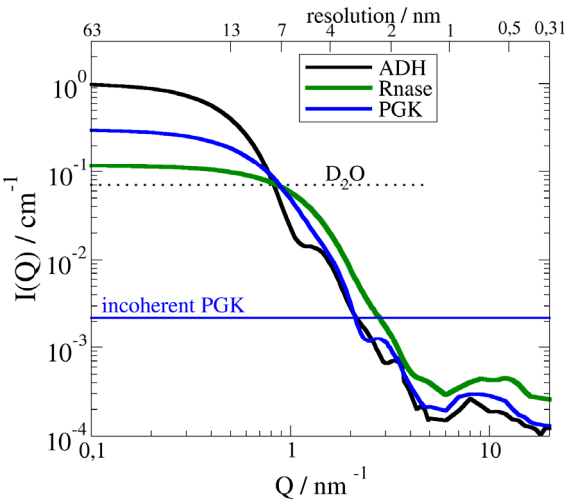


Figure 1 Neutron scattering of single proteins (ADH: alcohol dehydrogenase M_w=145kDa; PGK: Phosphoglycerate kinase M_w=44kDa; Rnase: Ribonuclease A M_w=14kDa) at a concentration of 10 mg/ml relative to the typical background of a D_2O buffer. The incoherent scattering is related to the exact amino acid composition, which determines the non-exchangeable hydrogen content responsible for the main contribution. The upper scale represents the spacial resolution as $2\pi/Q$.

arrangement in the protein. This large Q regime can only be analysed by polarisation analysis of polarized neutrons to extract coherent and incoherent contributions.
A main goal of the analysis of small angle scattering data is to extract a model of the protein in solution and how it is influenced e.g. by substrate binding.

2.2 Rigid body refinement

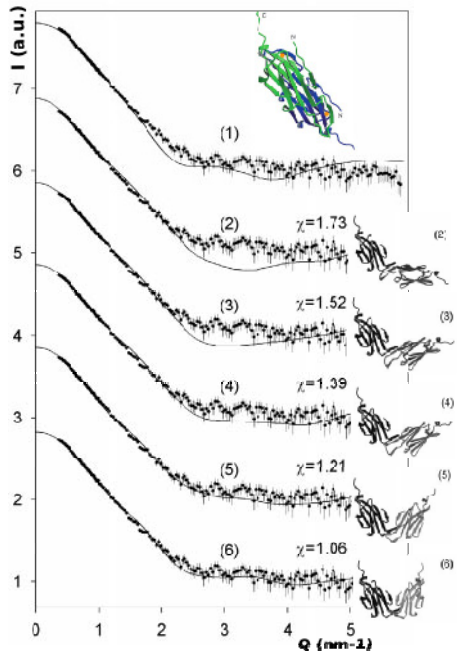


Figure 2 Experimental scattering data I(Q) compared to atomic model calculations with changed orientation between rigid monomers of the α-crystalin domain as indicated at the right side. (1) the crystallographic dimer of MjHSP16.5; (2-6) dimeric homology models with different orientation.

A common problem for protein complexes is the known structure of subunits without knowledge about the quaternary structure. A reason could be that the larger structure cannot be crystallised for X-ray crystallography and/or the complex is to large for NMR studies. Here the known subunits can be combined in a new way to find the quaternary structure taking into account known interfaces.
Feil et al. report about a truncation mutant of αB-crystalline, which is a heat shock protein and a major eye lens protein [10]. The mutant is a dimeric protein of the α-crystalin domain. A homology model is constructed based on the similarity to Methanococcus Jannaschii, which shares 23% of the monomeric sequence and 50% are chemically similar. This model consists of 2 monomers with mainly β- sheet structure in a flat geometry and both monomers lying flat on each other (see Figure 2 (1)). To check the validity of this model X-ray synchrotron measurements were analysed. Figure 2 shows the resulting scattering data with different model calculations based on the atomic structure. In (1) the homology model shows significant deviations compatible with a dimeric structure but not compatible with the flat arrangement. (2-6) show calculations based on generated models compatible with an connecting interface known from spin labelling studies and the rigid monomer structure. The final best description is shown as (6) with an open structure and an a V-shaped arrangement. This arrangement allows a higher flexibility and has a larger surface area compared to the compact homology model. Both, higher flexibility and larger surface area, may be important for the biological function as chaperon and for the formation of αB-crystallin complexes.

2.3 Bead modelling

If now prior knowledge about the structure of a protein is known, methods can be used that try to find compatible structures that describe the measured data. These algorithms cannot find a unique solution because of the invers scattering problem and the restricted Q range of typical small angle scattering data. Therefore multiple solutions are averaged to find the best compatible solution. The accessible maximum Q also defines a minimal length scale,

therefore details of smaller size cannot be resolved and suggest a coarse grained approach. While rigid body refinement uses larger known subunits as coarse grain elements, without such knowledge small spheres can be used as subunits.
A well-known set of programs is developed in the Svergun group [11]. Lipfert et al demonstrate the usage of the program DAMMIN to explore the structure of the VS Ribozyme [12]. Ribozymes are enzymes build from fragments of RNA, which catalyse chemical reactions. The VS ribozyme is a large ribozyme build from 5 helical sections without known tertiary structure that is able to cleave a phosphodiester bond in the RNA backbone.
DAMMIN uses a random distribution of spheres in an enclosing volume as a starting point as shown in Figure 3a to calculate the expected scattering. Random changes of the spheres by a simulated annealing algorithm change the configuration and are weighted by a looseness criterion and the quality of the description of the measured data. An animated example can be found in reference 13. Several solutions (Figure 3b) are aligned by the program DAMAVER and averaged to a model describing the occupancy of different positions (Figure 3c, occupancy coloured). Figure 3d shows the final reconstructed density map that is used by Lipfert et al. to reconstruct the final tertiary structure from the subunits as shown in Figure 3e. From analysis of the tertiary structure the position and orientation of the substrate in the VS ribozyme was determined.

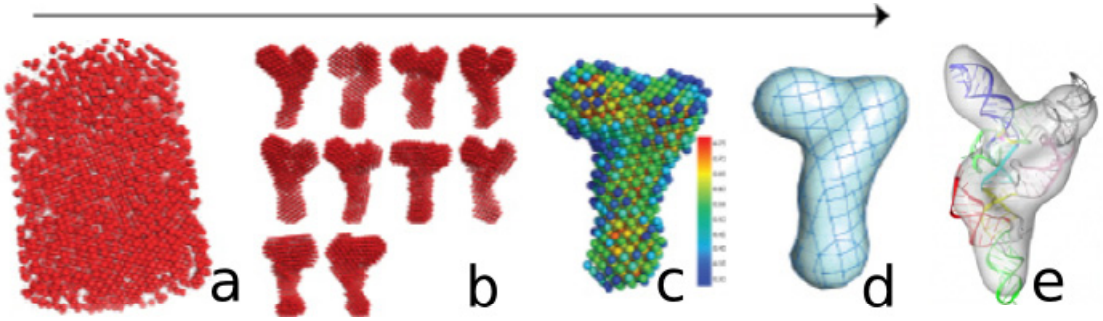


Figure 3 Shape reconstruction starting from a random distribution to the final solution. The tertiary structure is reconstructed from the subunits to fit to the final solution. a: starting condition, b: ten independent results, c: averaged results, d: density map, e: ribozyme structure fit in density map. Details with an animation of the annealing process can be found in supplementary information of Lipfert et al [12, 13]

2.4 Structure deformation by normal modes

The protein structure in the PDB is in most cases measured by X-ray crystallography and thus influenced by the geometrical restrictions due to the next neighbours in the crystal lattice. In solution the protein has more freedom to find a structure close to the minimum of the potential energy and the protein will adopt its conformation. Changing pH, salt concentration, temperature or pressure shift the minimum of the potential energy and the protein will adopt to these changes. In the same sense the binding of a substrate changes the energy of the complete system and may lead to a configuration change. The pathway, which is followed, will be a minimum energy pathway in the complex energy landscape of the protein. The energy of a configuration includes contributions from bonded interactions, long-range electrostatic and van der Waals forces. Additional contributions come from the water and salt atoms in the solution that modify the energy of the protein-water system. These are the interactions that are included in each step of a molecular dynamics simulation. For a given

configuration the potential can be approximated by a simplified harmonic force field, if a potential minimum is reached [14]. Within this approximation normal modes as solution of the 3N dimensional eigenvalue problem of $K \cdot e_i = \lambda_i \cdot e_i$ with the 3Nx3N force matrix K, 3N dimensional eigenvectors e_i and eigenvalues λ_i describe the basic eigenmotion of the N protein atoms. If we are interested only in large scale configurational changes a simplified approach can be used, which uses a coarse gained description, e.g. only the c-alpha atoms of the residues, with elastic springs between neighbouring elements. Such an elastic network model can be thought of as a rubber model of the protein to find the weak parts, where the protein can be moved with the weakest force. To modify the configuration of a protein the normal mode eigenvectors can be used as template for small structural changes along the normal mode vectors.

Inoue et al. use this method to elaborate the structural changes of Phosphoglycerate Kinase (PGK) due to the substrate binding [15]. PGK is a two domain protein with a hinge between the two main domains as shown in Figure 4. In previous work it was suggested by homology

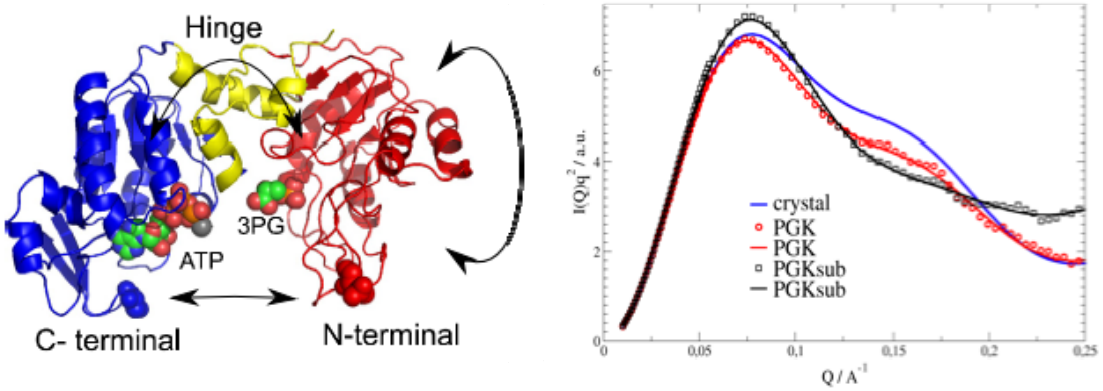


Figure 4 left: Image of the ternary 3PG-MgATP complex of yeast PGK (PDB code 3pgk) showing the N- and C-terminal domains (red and blue, respectively) and the hinge region (yellow). The arrows indicate the normal mode movement's: mode 7 torsion of N-terminal against C-terminal, mode 8 closing cleft movement and mode 9 perpendicular hinge movement to mode 8. The substrates are given as spheres. Right: Experimental formfactor of PGK extracted from SANS in a Kratky plot ($I(Q)Q^2$ vs. Q) with and without bound cofactor (PGKsub and PGK, respectively) compared to the crystal structure calculation (blue) and the refined models (red and black lines). The Kratky plot highlights smaller differences at larger Q values.

studies that the cleft with the active centre in between the main domains near to the hinge is closed due to the substrate binding of adenine triphosphate (ATP) and 3-phophoglycerate (3PG). This was also suggested by studies of the radius of gyration R_g, which showed a more compact configuration with bound substrate. The closing of the cleft is necessary for the transfer of a phosphate group from 1,3-phosphoglycerate to adenosine diphosphate to recharge it to ATP with a stabilisation of the phosphate group by a tripyramidal configuration. The 3 atoms contributing to this stabilisation are located on both main domains and approach close enough only in a closed configuration [16].

Inoue et al. investigated the change due to substrate binding in more detail by modelling the measured SANS data with a configuration changed by the first three normal modes from an elastic network analysis done with ElNemeo [17]. The first three elastic normal modes

represent in this case the basic movement of a two body system connected by a hinge: torsion and two perpendicular bending motions. Compared to the crystal structure PGK with and without bound substrate show a different scattering as shown in Figure 4 with a decrease of R_g from 23.7 Å to 22.5 Å upon substrate binding. The change from the crystal structure to the solution structure is mainly due to a mode 9 displacement (mean atomic displacements of -1.4 Å, 0.3 Å, 2.2 Å for modes 7–9, errors ~0.7 Å). The change due to the bound substrate is mainly along mode 8 and 9 closing the cleft (0.8 Å, -3.7 Å, -1.9 Å). The configurational change can be related to a functional relevance because the normal modes describe also the configurational change in the active center. Here it is found that the three atoms stabilizing the phosphate group are not close enough to allow the transfer of the phosphate group if the substrates are bound. Thus a detailed analysis confirms the closing of the cleft, but shows that the closing is not large enough for protein function.

3 Protein dynamics

Protein dynamics involves different processes at different length- and timescales [18]. The relative vibration of bonded atoms as hydrogen bonded to the backbone nitrogen or atom groups (e.g. side chain methyl groups) have spatial extends of several angstroms with pm amplitudes. Local atomic fluctuations have contributions from elastic vibrations and sidechain movements as the torsion of tyrosine sidechains. At the protein surface loops can move freely. On larger length scale we find cooperative motions like hinge motions or torsions between more rigid domains. On top of these internal motions we have translational and rotational motion of the complete protein. Especially translational diffusion is dependent on the surrounding due to geometrical restrictions by obstacles or for concentrated protein solutions by the presence of the next neighbouring proteins.

In general all these motions are dependent on each other and are coupled. The local atomic fluctuations lubricate the domain motions on larger length scales and the domain motions change the shape of the protein and thus influence the friction with the solvent for rotation and translation, if the size is increased. Rotation and translation are in general coupled, e.g the rotation of a screw will result in a translation along the rotation axis. For simple shaped bodies the coupling is very weak or zero for symmetric shapes. Nevertheless it is a good working approach to separate different contributions as being independent.

Coherent and incoherent scattering give different view on the same dynamic processes. Incoherent scattering shows us the atomic self correlation because of the missing interference between different atoms. Because specific atoms are part of larger complexes like a sidechain or the protein itself they are incorporated in all collective motions, which also show up in the incoherent scattering beside e.g. atomic vibrations. Coherent scattering has a strong contribution of interatomic interferences, which for small Q leads to the dominance of the coherent scattering as shown in Figure 1. Consequently the correlated domain motions on larger scale as shape deformations or diffusion dominate coherent scattering dynamics at low Q. Dependent on both absolute contributions at a specific Q we always measure a mixture of coherent and incoherent dynamics. On the other hand the dynamic timescale differ dependent on the observed process. If we look at atomic vibrations on sub picosecond scale the rotational and translational diffusion can be neglected because diffusion is to slow to be seen within the observation time window. On the other side of diffusion dominated measurements on nanosecond scale the picosecond dynamics of vibrations averages out and we observe an

average position of the atoms. For measurements on intermediate timescale we have to include fast and slow contributions.

3.1 Atomic vibrations

The fastest movements with highest energy are atomic vibrations as they are common for any molecule. Figure 5a shows as an example the geometry of a torsion around the C-C bond of a methyl group, a symmetric stretching of the C-H bonds and a symmetric bending of the c-H bonds of the sidechain of alanine. These are only a few possible vibrations of alanine and

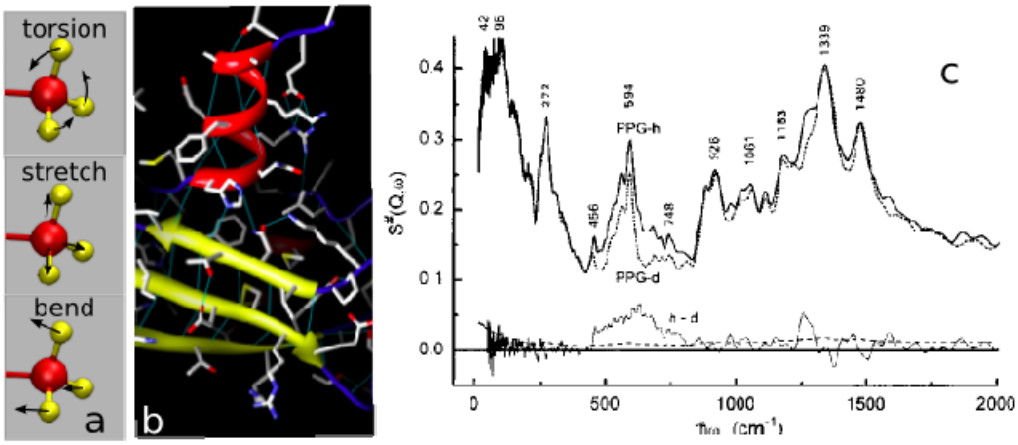


Figure 5 a: Example geometries of torsion, symmetric stretching or bending vibrations of a methyl group as the side group of an alanine residue. b: Secondary structure elements β-sheets and α-helices (in yellow, red) with sidechain elements. The secondary structure is stabilized by the hydrogen bonds (thin blue lines). c: Vibrational spectrum of synthetic polypeptide (PPG)n associated in a right-handed supercoiled triplehelical arrangement as found in natural collagen measured at the time-focused crystal analyzer spectrometer (TFXA) at ISIS, UK by Middendorf et al. [20] .

each amino acid has different vibrational frequencies dependent on the atomic structure. For an overview see Barth et al [19]. The exact frequency depends not only on the geometry as for free molecules of the same architecture, but also on the direct environment and neighbouring amino acids. Hydrogen bonds or polar interactions alter the vibration frequency. Figure 5c shows an example spectrum of $(PPG)_n$ an synthetic polypeptide with a similar structure to natural collagen [20]. The frequency range reaches from 100 cm^{-1} to about 2000 cm^{-1} (12.4 meV – 250 meV) describing motions on a timescale of 0.01ps to 0.3 ps. The difference in the spectra is related to a partly deuteration that changes the frequency and the amplitude of specific vibrations due to the change of the mass. In this way the specific exchange allows to separate some of the vibrational modes present in the sample. Typical instruments with the necessary large energy transfer are time of flight instruments.

An alternative approach to examine vibrational modes is infrared spectroscopy. For protein the Amide I absorption band between 1600 cm -1 and 1700 cm^{-1} is of importance because it allows the determination of the content of secondary structure in the protein. In this absorption band the C=O stretch vibration of the peptide backbone is dominant. As shown in Figure 5b the hydrogen bond between the C=O and the N-H of a different amino acid stabilise the secondary structure elements. A specific hydrogen bond modifies the stretching vibration

in a way that is characteristic for the local geometry defined by the secondary structure. Measuring the different contributions allows extracting the fractional content of secondary structure elements.

3.2 Local atomic fluctuations

At larger Q the incoherent dynamics dominates the neutron scattering in the incoherent intermediate scattering function S(Q,ω). For larger energy transfers up to 10 meV quasielastic scattering and inelastic scattering are measured at time of flight instruments as shown in Figure 6. At largest energies an inelastic contribution of structural vibrations like phonons in a disordered lattice is generally observed (DHO in in Figure 6). The quasielastic contributions to the signal result from local motions of the hydrogen atoms, which give the largest contribution to the measured signal. Pieper et al. compare the quasielastic dynamics of bacteriorhodopsin (BP), a light driven proton pump in the cell membrane of *Halobacterium salinarum* in a dark state and in a excited state, reached by the absorption of a short fs laser pulse [21]. The exited state is measured within a pump-probe setup by synchronising the laser pulse and the data acquisition of the scattered neutrons. During the absorption of light a structural change occurs that reorients the retinal molecule inside of the protein to be accessible from the cytoplasmic side of the cell allowing the proton transfer. Two Lorentzians related to free diffusion of the protons with relaxation times of 12 ps and 1 ps describe the quasielastic contribution. The excitation reduces the slower contribution and increases the faster contribution to the total signal. It was concluded that beside the configurational change of the structure also a modulation of protein dynamics occurs.

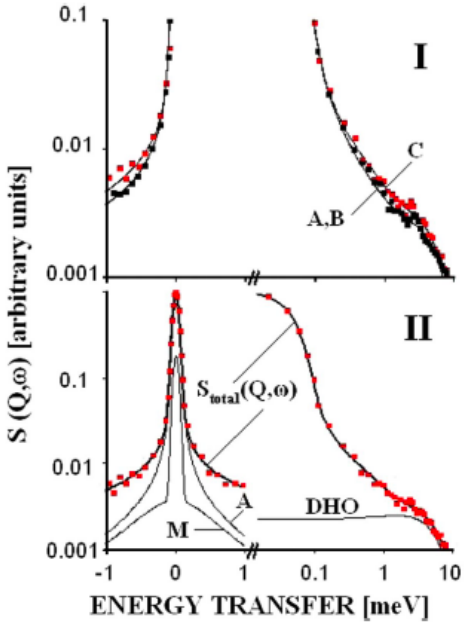


Figure 6 Quasielastic neutron scattering spectra measured at NEAT, HMI, Berlin with and energy resolution of 93μev at room temperature [21]. Data are averaged over all Q vectors resulting in $Q=1.5A^{-1}$. I: A, B measurements of bacteriorhodopsin in the dark, C (red) shortly after excitation with light. II: Contributions to the signal from the M intermediate, the groundstate A and an elastic contribution described by a damped harmonic oscillator (DHO).

If the resolution of the used instrument is decreased the description of the proton dynamics as a free diffusion for the short times breaks down because the observation timescale is long enough to show geometrical restrictions from the neighbouring atoms. Stadler et al. measured the dynamics of hemoglobin in red blood cells in vivo at different temperatures [22]. Hemoglobin is responsible for the oxygen transport in red blood cells and stores the oxygen bound to a heme group present in a pocket of the protein. The inelastic neutron scattering experiments of Stadler at al. measured at a time of flight instrument shown in Figure 7A, have a small contribution due to diffusion close to the resolution (black dashed line close to blue resolution). The accessible timescale reaches from the resolution limit of about 80 ps to about

5ps at the largest energy transfer. The dominating feature is the faster relaxation (red line). Fast and slow contributions are described by Lorentzian functions in S(Q,ω). The elastic incoherent scattering function (EISF) is the value at ω=0 of the incoherent intermediate scattering S(Q,ω=0) and contains information about the geometry of dynamical processes.

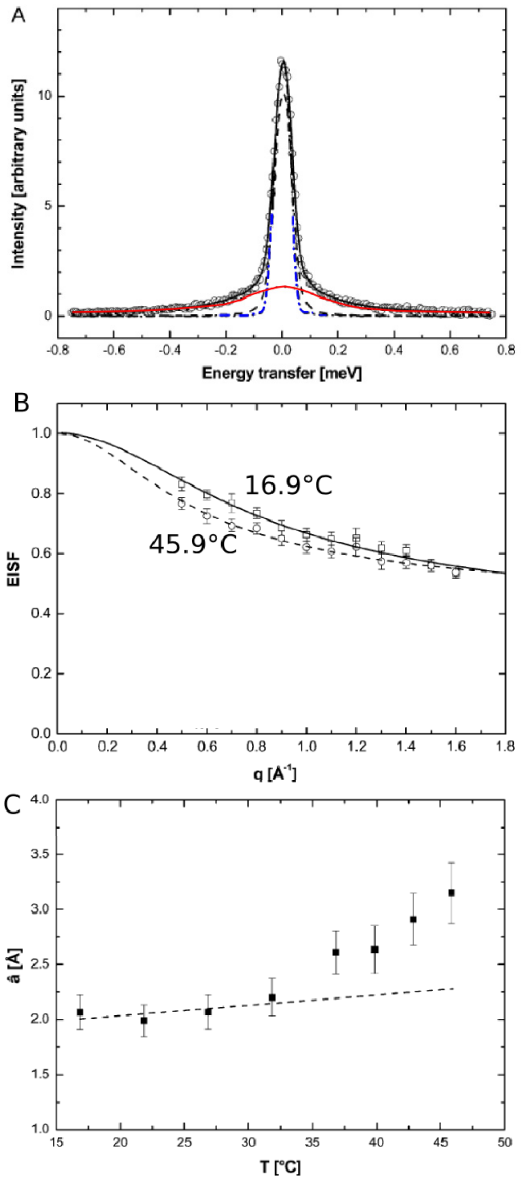


Figure 7 A: Quasielastic neutron scattering spectra of hemoglobin measured at FOCUS, SINQ, Swiss with a resolution around 50μeV (blue) at Q=1.6 A^{-1} [22]. A slow component (black dashed) and a faster component (red) are visible. B: EISF(Q) as the maximum amplitude extracted from the fast component in A. The lines correspond to a model of jump diffusion within a sphere of radius *â* (including a distribution of sizes) at indicated temperatures. C: Sphere radius *â* dependent on temperature.

Figure 7B shows the Q dependence of the EISF of the fast component extracted from measurements as shown in Figure 7A for two temperatures. The lines are model functions according to jump diffusion within a sphere with radius *â* assuming additionally a distribution of sphere sizes to describe the varying local structure. A reduction of the amplitudes at higher temperature is observed. For details see ref. 22. The restricting sphere for a hydrogen jump is the contour of the van de Waals radius of neighbouring atoms. If the observation time of the instrument is larger than the time a hydrogen needs to diffuse over a sphere radius, the restriction is visible. Stadler et al. observed an increase of the restricting sphere radius at body temperature of 37°C (see Figure 7C) which is correlated to a partial loss of α-helical content and a higher flexibility, as also observed by micropipette experiments. The observations are interpreted as a partial unfolding of the protein allowing the sidechains to move in a larger space and enabling larger amplitude motions.

A detailed analysis of the slower diffusion component, which is attributed to global diffusion of the protein, requires a better resolution of the instrument as shown in Figure 8. Stadler et al. measured hemoglobin dynamics with the backscattering instrument IN16, ILL at a resolution of 0.9μeV [23]. At larger Q the diffusional dynamics described by the slower Lorentzian comprises contributions from translational and rotational diffusion. Rotational diffusion increases the apparent diffusion coefficient relative to the expected translational diffusion coefficient for large Q by a factor of 1.27 (for an explanation see next section). The increased resolution allows the access of

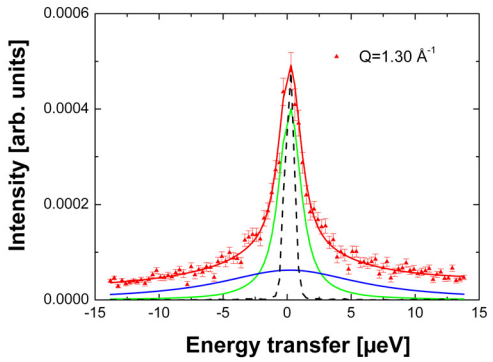


Figure 8 Quasielastic scattering of hemoglobin measured at IN16, ILL, France with a resolution of 0.9µeV (black dashed). The scattering is described by the sum (red) of a fast Lorentzian function (blue) describing the internal proton dynamics and a slower Lorentzian (green) describing the combined rotational and translational diffusion.

slower dynamic processes up to 4 ns compared to measurements with a resolution of 0.2 ns (resolution 17µeV) at IRIS, ISIS. Hemoglobin in red blood cells has an approximate volume fraction of 25%. In the short time measurements, the protein diffuses not far enough to feel the hindrance of the surrounding proteins, so the short time self diffusion is observed. For longer times the surrounding proteins hinder the protein diffusion and a reorganisation of the proteins is necessary to allow larger diffusion pathways. This rearrangement needs more time and slows down the overall diffusion as is observed [23].

3.3 Integral measurement of local dynamics

The elastic intensity of the EISF S(Q,) at Q=0 allows a quick access to general dynamical properties of a protein. The measurement is done by an elastic scan, which is much faster than a scan of all accessible energy transfers. Extrapolation of the Q dependence of the EISF similar to Figure 7B to Q=0 with $< u^2 >= -6 \lim \partial \ln(I_{el a}(Q)) / \partial Q^2$ results in the mean square displacement present in the sample [25]. An atom moving out of the Volume defined by $2\pi/Q$ within the time limit defined by the instrument resolution shows up contributions in the inelastic scattering. The limit $Q \rightarrow 0$ extends this to infinite displacements resulting in the mean square displacement within the resolution timewindow. It should be noted that the second limit still depends on the resolution of the used instrument.

Changing the environmental conditions changes $<u^2>$ e.g. drying the protein reduces the fluctuations and increasing temperature increases the fluctuations.

Figure 9a shows measurements of Malate dehydrogenase and Lactate dehydrogenase which have their maximum activity at 90°C and 37°C, respectivly [24]. From the slope of the

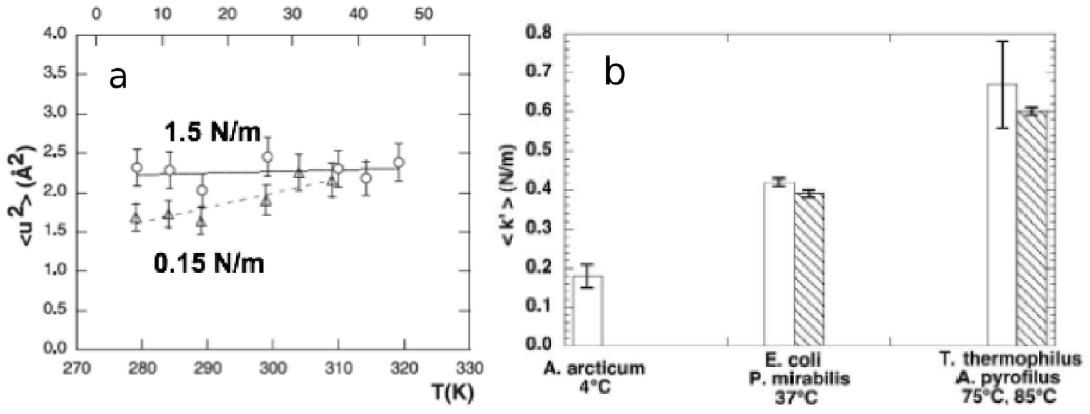


Figure 9 a: Mean square displacement $<u^2>$ of homologues Malate dehydrogenase (circles) and Lactate dehydrogenase (triangles) from neutron scattering at IN13 measured with a resolution of 8µeV (timescale 100ps) [24]. B: Mean force constant <k'> measured for different bacteria which live at ambient temperatures of 4°C, 37°C and above 75°C.[26]

temperature dependence an effective force constant k can be determined according to $\langle u^2 \rangle = kT/\langle k \rangle$ if an harmonic mean potential is assumed [25]. This effective force has no physical meaning, but it can be used to compare measuremnts under changing conditions on the same instrument with the same resolution. The result of Tehei et al. show for both proteins the mean square displacement are the same even if the effective force is one order of magnitude different. The same was observed in examinations of complete cells of bacteria which live at extreme temperature conditions of 4°C, 75°C, 85°C compared to 37°C [26]. In both cases the composition of aminoacids seems to be addapted to the environmental conditions to allow a flexibility neccessary for biological function. For larger temperatures this requires a stiffer protein structure withstanding the larger thermal energy at larger temperatures.

3.4 Large scale domain motions

Inelastic neutron scattering experiments on proteins are usually performed at high Q where the incoherent scattering dominates and atomic fluctuations on a picosecond timescale are measured. At low Q the coherent scattering dominates and in general slower large scale collective dynamics in the protein can be observed. Inelastic neutron scattering techniques usually measure the Fourier transform of the space-time correlation function S(Q,ω). The neutron spin-echo technique implicitly performs a cos-Fourier transformation back to (Fourier-) time and therefore yields the intermediate scattering function $I(Q,t)$ in the time domain together with a high resolution in time [27].

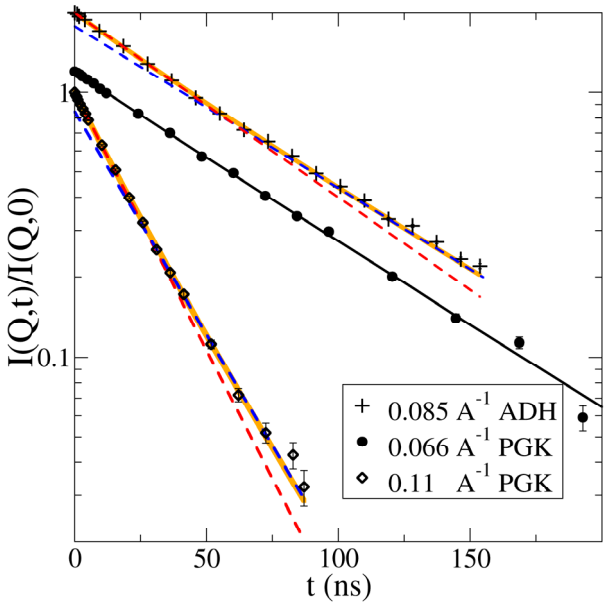


Figure 10 Intermediate scattering functions for selected Q values of ADH and PGK with substrate for selected Q. The red broken lines are the initial slope extrapolated to long times. The blue broken lines are the long time limit as expected for $D_0(Q)$ including corrections by the structure factor and the hydrodynamic function extrapolated to t=0. For PGK Q=0.066 A^{-1} is shifted by a factor 1.2, for ADH Q=0.067 A^{-1} is shifted by a factor 2 for clarity.

Typical measurements of the intermediate scattering function as measured by NSE are shown in Figure 10 as examples [28]. Phosphoglycerate kinase (PGK) is a protein involved in glycolytic pathway with a molecular weight of 44500 Da. It has 2 main domains connected by a hinge allowing fluctuations between the subdomains. Yeast alcohol dehydrogenase (ADH) with a molecular weight of 147kDa converts ethanol to acetaldehyde, a reaction that can be catalysed in both directions. ADH is a tetramer with each monomer having a dumbbell like structure with two globular domains separated by a cleft. The spanned Q for coherent scattering ranges from 0.02 $Å^{-1}$ to about 0.25 $Å^{-1}$ where the coherent signal for concentrations of 50 mg/ml equals the D_2O background. The initial slope of the intermediate scattering function for $t \rightarrow 0$ comprises the

dynamics due to the translational and rotational diffusion of the protein together with any additional internal dynamics of the protein.

The dynamics of proteins in solution is dominated by the translational and rotational diffusion of the protein. The single protein diffusion may be approximated by the diffusion of a non spherical rigid particle at finite dilution. The observable single protein diffusion $D_0(Q)$ at infinite dilution for a neutron scattering experiment is given by [28].

$$D_0(Q) = \frac{1}{Q^2 F(Q)} \sum_{j,k} \left\langle b_j \exp(-i\mathbf{Q}\mathbf{r}_j) \begin{pmatrix} \mathbf{Q} \\ \mathbf{Q}\times\mathbf{r}_j \end{pmatrix} \begin{pmatrix} D_{TT} & D_{TR} \\ D_{RT} & D_{RR} \end{pmatrix} \begin{pmatrix} \mathbf{Q} \\ \mathbf{Q}\times\mathbf{r}_k \end{pmatrix} b_k \exp(i\mathbf{Q}\mathbf{r}_k) \right\rangle$$

with the 6x6 diffusion tensor D comprising translational and rotational contributions of the rigid protein. The main contributions arise from the diagonal elements of translational diffusion D_{TT} and rotational diffusion D_{RR} and correspond to the scalar diffusion coefficients as the average over the corresponding trace.

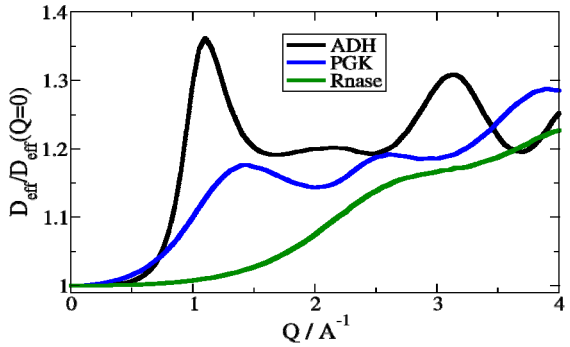


Figure 11 $D_0(Q)/ D_0(Q=0)$ for ADH, PGK and Ribonuclease, a small protein of molecular weight 13700Da. The diffusion coefficient $D_0(Q=0)$ at a temperature of 10°C according to the HYDROPRO result is 2.7 A^2/ns, 3.94 A^2/ns and 6.3 A^2/ns for ADH, PGK and Rnase, respectively.

By evaluating the hydrodynamic friction of an equivalently shaped object D can be calculated by the computer code HYDROPRO [29] for a given protein PDB structure. The resulting $D_0(Q)/D_0(Q=0)$ is shown in Figure 11. At large length scales (small Q) the rigid body is seen as shapeless, quasi point-like object, showing only the constant Q independent translational diffusion D_{T0} as observed by dynamic light scattering (DLS) at infinite dilution. The time dependent intermediate scattering function is a single exponential as seen in Figure 10 for lower Q. The additional contribution from rotational diffusion leads to an increase in the effective rigid body diffusion $D_0(Q)$ above D_{T0} proportional to the rotational diffusion. It is observed, if the investigated length scale ($\sim 2\pi/Q$) is smaller than a length typical for the deviation of the protein shape from spherical symmetry. At larger Q the already mentioned increase of an apparent diffusion coefficient due to rotational diffusion to about $1.27D_0$ is observed.

At finite concentrations interparticle interactions cannot be neglected, and the collective translational diffusion coefficient is $D_T(Q)=D_{T0}H_T(Q)/S(Q)$, where $H_T(Q)$ is the hydrodynamic function representing hydrodynamic interparticle interactions mediated by the solvent and the structure factor S(Q) representing the direct interactions as measured by SANS. D_{T0} is the Q independent single particle translational diffusion. The hydrodynamics influences also the rotational diffusion D_R but independent from Q by $D_R=D_{R0}H_R$ with the single particle rotational diffusion constant $D_{R0}(30)$. All of these effects influence the observed translational and rotational diffusion also at large Q accessed by the incoherent scattering in previous sections. For lower concentrations below 10% volume fraction the structure factor S(Q) equals one at larger Q and $D_T(Q)$ is only slowed down by hydrodynamic interactions.

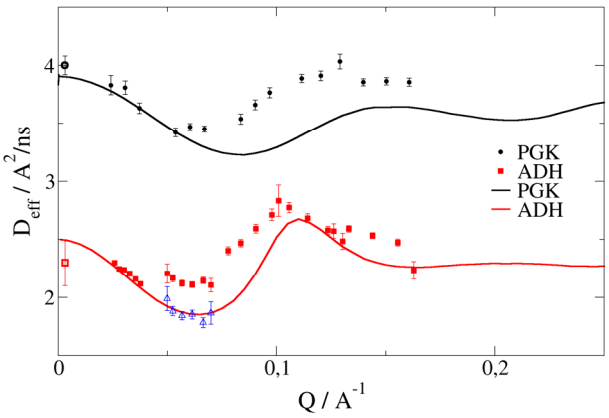


Figure 12 Effective diffusion as result of o second order cumulant fit for PGK at a temperature of 10°C (black) and ADH at 5°C (red). The blue points show the effective diffusion as a result of a single exponential fit for times t> 60 ns as the long time limit for ADH. For both proteins the first point at Q=0.003 A-1 represent DLS measurement, which resulted in concentration independent diffusion constants below 5% w/v. The lines are expectations for the effective diffusion of a rigid protein including effects of the structure factor and the hydrodynamic function.

Figure 12 shows the results for PGK and ADH in comparison to the results of dynamic light scattering experiments (as the limit for small Q) under the same conditions (for experimental details on ADH see ref. 28). For both proteins we find a decrease at low Q values followed by an increase at higher Q. The lines represent the expectation for the rigid protein according to the results of HYDROPRO including the effect of the structure factor S(Q) on the translational diffusion, which was measured independently by SANS, and the effect of the hydrodynamic function H. Already the modification due to the structure factor S(Q) results in a good description of the low Q decrease in $D_{eff}(Q)$ under the assumption of a constant $H_T(Q)$. H_T values of 0.81 for ADH and 0.74 for PGK were found. The change in the effective diffusion is due to additional internal dynamics.

From the comparison of the expected behaviour for a rigid protein one can conclude that large scale fluctuations of entire domains are mostly visible when also the rotational diffusion starts to contribute to the measured initial slope. This can be understood as large domain fluctuations mainly change the shape of the protein and therefore have the same length scale as shape fluctuations seen by the scattering vector Q due to rotations.

To extract these internal dynamics from the measured intermediate scattering function the decomposition of rotational diffusion into spherical harmonics by Lindsay [31] with scalar translational and rotational diffusion constants D_{T0} and D_{R0} of the rigid protein to describe the time dependent diffusion contribution can be used. Additional internal dynamics is described by an additional exponential relaxation with a common relaxation time and Q dependent amplitude A(Q).

From this direct comparison for PGK a relaxation time of about 30ns for PGK with substrate and additional amplitude of about 0.12 for the high Q data in Figure 10 is evaluated. Without substrate the relaxation time is 60 ns. Analysis of all available Q data allows the determination of the Q dependent amplitude A(Q) of the specific internal protein dynamics as shown for PGK with and without substrate in Figure 13. To describe the effect of specific protein movements onto the intermediate scattering function the occurring large scale domain motions are described in a first order approximation for small displacements [32] by elastic normal modes [15].

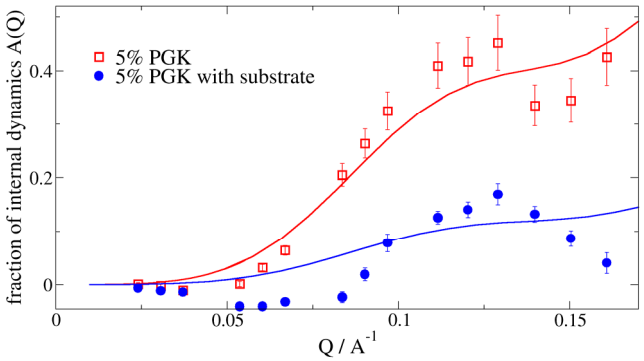


Figure 13 Fraction A(Q) of the internal dynamics to the intermediate scattering function for PGK with substrate (blue, relaxation time 30 ns) and without substrate (red, relaxation time 60 ns). The first three normal modes display the simplest relative motions of a two body system with a flexible hinge, which are torsion and 2 perpendicular hinge motions as shown in Figure 4left. A model allowing fluctuations along these modes individually is shown as lines.

The resulting contribution to the intermediate scattering function for PGK is shown in Figure 13 in comparison to the experimental results. As result it was found that for PGK the bound substrate reduces the relaxation time from 60 ns to 30 ns in combination with a reduction of the amplitude by a factor of two. As shown in a previous section for PGK the configurational change due to substrate binding was not sufficient to reach the active configuration. The thermal domain dynamics with a 30 ns relaxation time has an amplitude that allows to reach the active configuration. Here the dynamics drives the configurational change to allow biological activity different to older models.

4 Summary

The structure and dynamics of proteins are of importance for the biological function in cells. Beside the 3dim atomic structure the changes of the structure due to environmental conditions or the biochemical partners in the active state are of general interest to understand the way proteins work. The large scale structure is examined by SAS methods, but needs a high quality interpretation to extract more than just the change of size. Experiments accessing the dynamics on timescales from sub picoseconds to hundreds of nanoseconds show neutron scattering techniques as a powerful tool to access fundamental dynamical processes.

1 http://www.wwpdb.org/

2 Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions *Nat Rev Mol Cell Biol,* 2005, 6, 197-208

3 Gunasekaran, K.; Tsai, C.-J.; Kumar, S.; Zanuy, D. & Nussinov, R. Extended disordered proteins: targeting function with less scaffold *Trends in Biochemical Sciences,* 2003, 28, 81-85

4 Smock, R. G. & Gierasch, L. M. Sending Signals Dynamically *Science,* 2009, 324, 198-203

5 Fischer E., Einfluss der Configuration auf die Wirkung der Enzyme, Ber. Dt. Chem. Ges. 1894, 27, 2985–93.

6 Koshland D. E., Application of a Theory of Enzyme Specificity to Protein Synthesis Proc. Natl. Acad. Sci. (1958) 44 (2): 98–104.

7 Nelson, D.L., Cox, M.M. Lehninger Principles of Biochemistry(5th ed). New York: W.H. Freeman and Company, 2008. pp. 96-100

8 Edman, P.; Högfeldt, Erik; Sillén, Lars Gunnar; Kinell, Per-Olof, Method for determination of the amino acid sequence in peptides". Acta Chem. Scand. 1950, 4, 283–293.

9 The infrared absorption of amino acid side chains Barth et al.,Progress in Biophysics & Molecular Biology 2000, 74,141

10 Feil, I. K.; Malfois, M.; Hendle, J.; van der Zandt, H. & Svergun, D. I. A Novel Quaternary Structure of the Dimeric Œ±-Crystallin Domain with Chaperone-like Activity *Journal of Biological Chemistry,* 2001, 276, 12024-12029

11 http://www.embl-hamburg.de/biosaxs/software.html
12 Lipfert, J.; Ouellet, J.; Norman, D. G.; Doniach, S. & Lilley, D. M.The Complete VS Ribozyme in Solution Studied by Small-Angle X-Ray Scattering*Structure,* 2008, 16, 1357-1367
13 http://www.dundee.ac.uk/biocentre/nasg/VS/SAXS_VS/1D_to_3D/LowerJunction.html
14 Tirion Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Phys Rev Lett,* 1996, 77, 1905-1908
15 Inoue R., Biehl R., Rosenkranz T.,Fitter J., Monkenbusch M., Radulescu A., Farago B., Richter D. Large Domain Fluctuations on 50-ns Timescale Enable Catalytic Activity in Phosphoglycerate Kinase Biophysical Journal, 2010, 99, 2309 - 2317
16 Bernstein, B. E., P. A. Michels, and W. G. Hol. Synergistic effects of substrate-induced changes in phosphoglycerate kinase activation. Nature, 1997, 385, 275–278.
17 http://www.igs.cnrs-mrs.fr/elnemo/
18 McCammon, J. A., Protein dynamics, *Reports on Progress in Physics,* 1984, 47, 1
19 Barth, A. The infrared absorption of amino acid side chains *Progress in Biophysics and Molecular Biology, Elsevier,* 2000, 74, 141-173
20 Middendorf, H.; Hayward, R.; Parker, S.; Bradshaw, J. & Miller, A. Vibrational neutron spectroscopy of collagen and model polypeptides *Biophysical Journal,* 1995, 69, 660-673
21 Pieper, J.; Buchsteiner, A.; Dencher, N. A.; Lechner, R. E. & Hau√ü, T. Transient Protein Softening during the Working Cycle of a Molecular Machine *Phys. Rev. Lett., American Physical Society,* 2008, 100, 228103
22 Stadler, A.; Digel, I.; Artmann, G.; Embs, J.; Zaccai, G. & B¸ldt, G. Hemoglobin Dynamics in Red Blood Cells: Correlation to Body Temperature *Biophys J, Cell Press,* 2008, 95, 5449-5461
23 Stadler, A. M.; van Eijck, L.; Demmel, F. & Artmann, G. Macromolecular dynamics in red blood cells investigated using neutron spectroscopy *Journal of The Royal Society Interface,* 2011, 8, 590-600
24 Tehei, M.; Madern, D.; Franzetti, B. & Zaccai, G. Neutron Scattering Reveals the Dynamic Basis of Protein Adaptation to Extreme Temperature *Journal of Biological Chemistry,* 2005, 280, 40974-40979
25 Zaccai, G. Neutron scattering perspectives for protein dynamics*Journal of Non-Crystalline Solids, 6th International Discussion Meeting on Relaxation in Complex Systems,* 2011, 357, 615-621
26 Tehei, M.; Franzetti, B.; Madern, D.; Ginzburg, M.; Ginzburg, B. Z.; Giudici-Orticoni, M.-T.; Bruschi, M. & Zaccai, G. Adaptation to extreme environments: macromolecular dynamics in bacteria compared in vivo by neutron scattering *EMBO Rep, European Molecular Biology Organization,* 2004, 5, 66-70
27 M. Monkenbusch, M., R. Schätzler and D. Richter, Nucl. Instr. and Meth. in Phys. Rex A., 1997, 399, 301
28 Biehl, R.; Monkenbusch, M. & Richter, D. Exploring internal protein dynamics by neutron spin echo spectroscopy *Soft Matter, The Royal Society of Chemistry,* 2011, 7, 1299-1307
29 J. G. de la Torre, M. L Huertas, and B. Carrasco, Biophys. J., 2000, 78, 719.
30 V. Degiorgio, R. Piazza and R. B. Jones, Phys. Rev. E, 1995, 52,2707.
31 H. M. Lindsay, R. Klein, D. A. Weitz, M. Y. Lin and P. Meakin, Phys. Rev. A., 1988, 38, 2614.
32 A. C. Zemach and R. J. Glauber, Phys. Rev,1956, 101, 118.

E 9 Glass Transition

U. Buchenau
Institut für Festkörperforschung
Forschungszentrum Jülich GmbH

Contents

[0]Lecture Notes of the 43rd IFF Spring School “Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources” (Forschungszentrum Jülich, 2012).

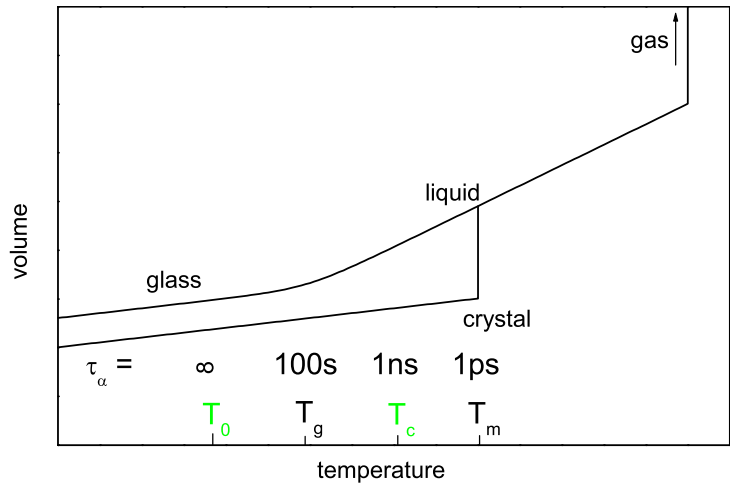


Fig. 1: *Volume-temperature diagram of the crystalline, glassy and liquid phase at ambient pressure.*

1 Introduction

The glass transition (the freezing of an undercooled and highly viscous liquid into a glass) is considered to be one of the great unsolved riddles of solid state physics. The main riddle is the atomic mechanism of the flow process in the highly viscous liquid and its strong temperature dependence shortly before freezing. The freezing itself is recognized as a falling out of thermal equilibrium, a kinetic process. But it is not clear to which extent this kinetic process reflects a true thermodynamic second order transition at a lower temperature, the Kauzmann or Vogel-Fulcher temperature. The present paper describes the situation of our present knowledge (or lack of knowledge) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

Liquids do not crystallize immediately after cooling below their melting points, because the crystallization requires the formation of crystal nuclei, a process which takes time. If this time is long enough to measure the properties of the undercooled liquid at all temperatures, one calls the liquid a glass former. Examples are vitreous silica (one mixes silica with metal oxydes to make window glasses), boron trioxyde, glycerol and selenium. Many polymers are good glass formers, like polystyrene, polycarbonate and polyisoprene (rubber), but others form a mixture of crystalline and amorphous domains like polyethylene.

At the glass transition temperature T_g, the undercooled liquid freezes into a glass, a solid with a nonzero shear rigidity. In many liquids, this glass temperature is about 2/3 of the melting temperature. From numerical simulation results (to be discussed in more detail in Section 4 on new developments), we now believe that every liquid can be cooled into a glass, provided the cooling rate is high enough. However, in real liquid water it is impossible to carry the heat fast enough away, so it can only be frozen into the glass state on the computer.

In undercooled liquids, the viscosity increases drastically with decreasing temperature. This increase reflects the dramatic increase of the structural relaxation time τ_α of the liquid (it is denoted by α to distinguish it from possible secondary relaxations β, γ etc. at lower temperatures in the glass phase). The structural α-relaxation is visible not only in mechanical, but also in dielectric, light scattering or heat capacity spectra. The glass transition occurs for a relaxation time of about 100 seconds.

Fig. 1 shows the cooling process from the gas phase into the liquid phase and then into the glass

phase in a volume-temperature diagram at ambient pressure. There is a large volume jump between gas and liquid, much larger than the volume jump at the first order phase transition between liquid and crystal at T_m. If the crystallization is avoided by fast enough cooling, the system enters the undercooled liquid and experiences the drastic slowing down of the relaxation time τ_α with decreasing temperature indicated in the figure. The glass temperature T_g depends on the cooling rate R according to

$$\tau_\alpha(T_g) \approx T_g/R, \tag{1}$$

because T_g/R is roughly the time needed to cool the whole energy of the sample away. So the faster the cooling, the higher is T_g. For a glass temperature of 200 K, a cooling rate of 1 K/s implies a τ_α of 200 s. As will be seen in the next section, τ_α is only a mean of a broad distribution of relaxation times, which makes the glass transition even more diffuse in temperature.

The thermal expansion of the glass is similar to the one of the crystal. It is due to the anharmonicity of the vibrations [1]. The thermal expansion of the liquid phase is a factor of two to four higher than the one of the glass phase (there is one important exception, vitreous silica, which will be discussed below).

From the theoretical understanding, there are two other important temperatures shown in Fig. 1. The first is the Vogel-Fulcher temperature T_0 below T_g where the viscosity and τ_α extrapolate to infinity. As will be seen, the Vogel-Fulcher temperature lies close to the Kauzmann temperature T_K, where the structural entropy of the liquid extrapolates to zero. This will be discussed in more detail in the next two sections. The second theoretical temperature, T_c, lies between T_g and T_m and is the critical temperature of the mode coupling theory [3]. The mode coupling theory is the most advanced liquid theory up to day. According to this theory, the viscosity should diverge at the critical temperature T_c. Therefore one has to invoke additional thermally activated hopping processes to explain why the real viscosity is still small at this temperature. A very recent theoretical treatment in terms of the replica technique [15] shows that T_c is the point in temperature where it becomes possible to measure a limiting shear modulus G_{high} for high enough frequency (G_{high} is sometimes denoted by G_∞, but this term should be reserved for the shear modulus of an instantaneous affine shear deformation [8]). From an empirical energy landscape point of view [5], T_c is the temperature below which one can distinguish between the fast picosecond relaxation processes and the slow thermally activated α-relaxation of the flow process close to the Maxwell time τ_M

$$\tau_\alpha \approx \tau_M = \frac{\eta}{G_{high}}, \tag{2}$$

where η is the viscosity. Since G_{high} is of the order of GPa, a viscosity of 10^{12} Pa s implies a Maxwell time of 1000 s.

The scattering methods discussed in the present spring school are of central importance for the study of the glass transition. Naturally, their time resolution is not good enough to study the α-process itself close to T_g, but they supply other essential information. This is illustrated in Fig. 2, which shows the mean square displacement of crystalline, glassy and liquid selenium as a function of temperature. There is a close parallel to Fig. 1, because the mean square displacement of the undercooled liquid, together with the anharmonicity of the interatomic potential, supplies the physical reason for the thermal expansion [1]. One realizes immediately that the understanding of the undercooled liquid requires a study of the atomic motion on the fast picosecond scale.

In a crystal, the simplest approximation relates the mean square displacement to a mean atomic frequency, the Debye frequency ω_D. If the temperature is high enough, one can use the classical

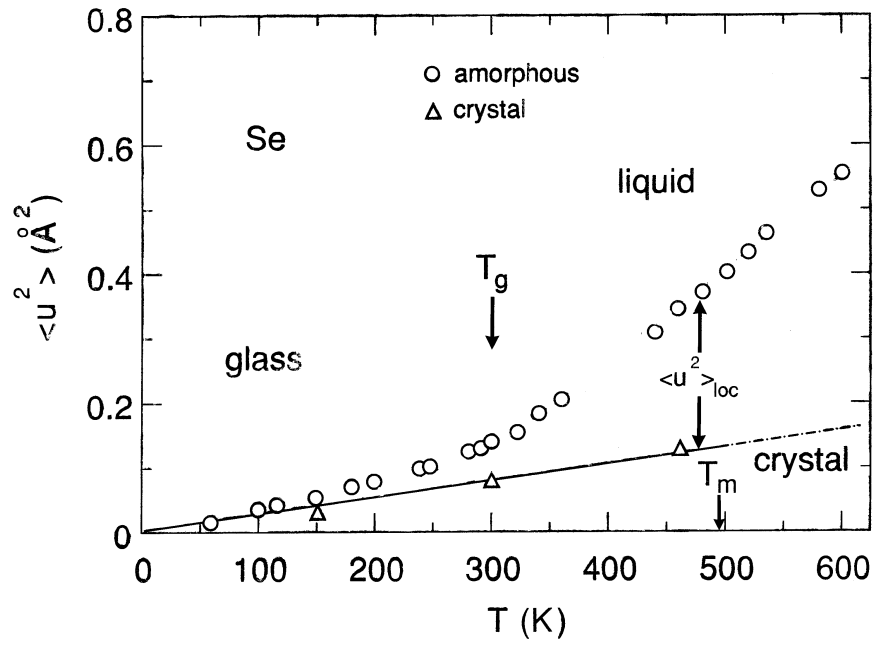


Fig. 2: *Mean square atomic displacements in crystalline, glassy and liquid selenium [16].*

approximation

$$< u^2 >= \frac{3k_BT}{M\omega_D^2}, \tag{3}$$

where M is the average atomic mass. In the simplest case of elastic isotropy, the Debye frequency is determined by the density ρ and the longitudinal and transverse sound velocities v_l and v_t, respectively

$$\omega_D^3 = \frac{18\pi^2\rho}{M(1/\mathrm{v}_l^3 + 2/\mathrm{v}_t^3)}. \tag{4}$$

In turn, the longitudinal and transverse sound velocities are related to the two elastic constants of the isotropic medium, the bulk modulus B_{high} and the shear modulus G_{high} by

$$\rho\mathrm{v}_l^2 = B_{high} + \frac{4}{3}G_{high} \qquad \rho\mathrm{v}_t^2 = G_{high}, \tag{5}$$

where the index $high$ indicates the elastic constant for high frequency and is not necessary in a solid. In a liquid, one has to make the distinction, because even the bulk modulus is markedly higher at high frequency.

The question is whether the simple description in terms of elastic constants still makes sense in a complicated glass former. This question can be answered by inelastic neutron scattering experiments, which are able to measure the vibrational density of states and its temperature dependence. It is this temperature dependence which is responsible for the strong rise of $< u^2 >$ in Fig. 2. It has been measured by Wuttke et al [17] in glycerol, a molecular glass former which exhibits the same $< u^2 >$-behavior as selenium and where the temperature dependence of the high-frequency elastic constants has also been measured [18]. The temperature dependence of the neutron spectra is shown in Fig. 3. The spectra are evaluated in terms of the vibrational

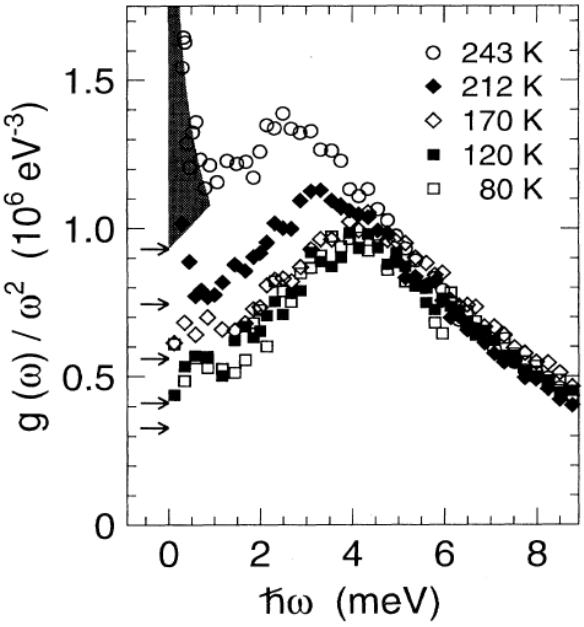


Fig. 3: *Measured spectra [17] in glycerol above and below its glass temperature T_g = 190 K, plotted as $g(\omega)/\omega^2$. The measured curves extrapolate to the arrows denoting the Debye expectation at the different temperatures. At 243 K, one begins to see the tail of the flow or α-process.*

density of states $g(\omega)$ divided by the frequency squared. In this plot, the Debye density of states

$$g_{Debye}(\omega) = \frac{3\omega^2}{\omega_D^3} \tag{6}$$

is a constant which can be calculated from the sound velocities at the given temperature. What one finds, is not a constant, but a broad peak sitting on top of the expected Debye density of states, the so-called boson peak, found universally in all glasses.

The boson peak is another one of the unsolved riddles of disordered matter; it is not clear whether it is due to resonant modes related to the low temperature tunneling states in glasses [19] (and to the plastic modes responsible for the shear thinning [20]) or simply to the force constant disorder in the glass [21]. This peak remains reasonably harmonic in the glass phase (not entirely; one can see some anharmonicity even below T_g), but begins to grow and shift to lower frequency in the undercooled liquid. At 243 K, one begins to see the high frequency tail of the flow process, because it begins to enter the nanosecond range accessible to neutrons, but at lower temperatures the undercooled liquid looks like a hot anharmonic glass in the neutron spectra.

Since the mean square displacement of the atoms is the area under the curve, it begins to grow markedly stronger than proportional to the temperature above T_g, as in Fig. 2. The anharmonicity of the potential is the same in liquid and glass. Therefore the stronger increase of the mean square displacement in the liquid with increasing temperature leads to a stronger thermal expansion [1]. Naturally, the fast vibrations supply only a part (if one analyses it quantitatively [22], only a smaller part, of the order of one quarter) of the additional thermal expansion or of the additional heat capacity in the undercooled liquid; the larger part comes from the slow fluctuations of the flow process at τ_α.

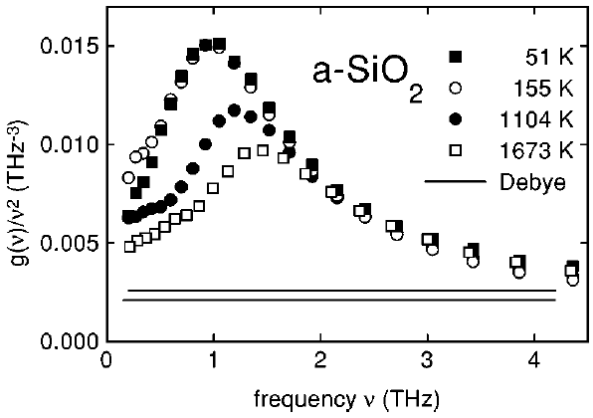


Fig. 4: *Measured spectra [23] in vitreous silica above and below its glass temperature T_g = 1473 K, plotted as $g(\omega)/\omega^2$. The upper line denotes the Debye expectation at 51 K, the lower line at 1673 K.*

As already mentioned above, silica shows an exceptionally different behavior. Fig. 4 shows the measured spectra [23] above and below the glass temperature of 1473 K in silica. One does indeed get the opposite behavior to the normal case in Fig. 3. Instead of growing and lowering its peak frequency, the boson peak decreases and goes to higher frequency with increasing temperature. In fact, one finds a negative thermal expansion below 150 K, where the boson peak vibrations dominate. At higher temperatures, one gets a small positive thermal expansion, because the vibrations at higher frequency have a positive Grüneisen parameter, the normal case [1]. The thermal expansion remains very small up to temperatures above T_g, making silica the closest example for a completely harmonic glass former that has been found so far. We will come back to this point in Section 4.

2 Kinetics of the glass transition

There are two characteristic properties of the kinetics of the α-process (the flow process) in undercooled liquids, the fragility and the stretching. The fragility characterizes the temperature dependence of τ_α and the stretching characterizes the width of the relaxation time distribution.

2.1 Fragility

The usual measure of the fragility of a glass former is the logarithmic slope of the relaxation time τ_α of the flow process

$$m = d\log\tau_\alpha/d(T_g/T)|_{T_g}, \tag{7}$$

in the so-called Angell plot [25] (see Fig. 5) of $\log\tau_\alpha$ as a function of T/T_g at $T/T_g = 1$.
It is useful to relate τ_α to an energy barrier V_α via the Arrhenius relation

$$\tau_\alpha = \tau_0 \exp(V_\alpha/k_B T), \tag{8}$$

where the microscopic attempt frequency is at 10^{-13} s, sixteen decades faster than the flow process at the glass temperature. The *fragility index* I is defined [10] by the logarithmic derivative $I = -d\ln V_\alpha/d\ln T$, taken at T_g. Then

$$m = 16(I+1), \tag{9}$$

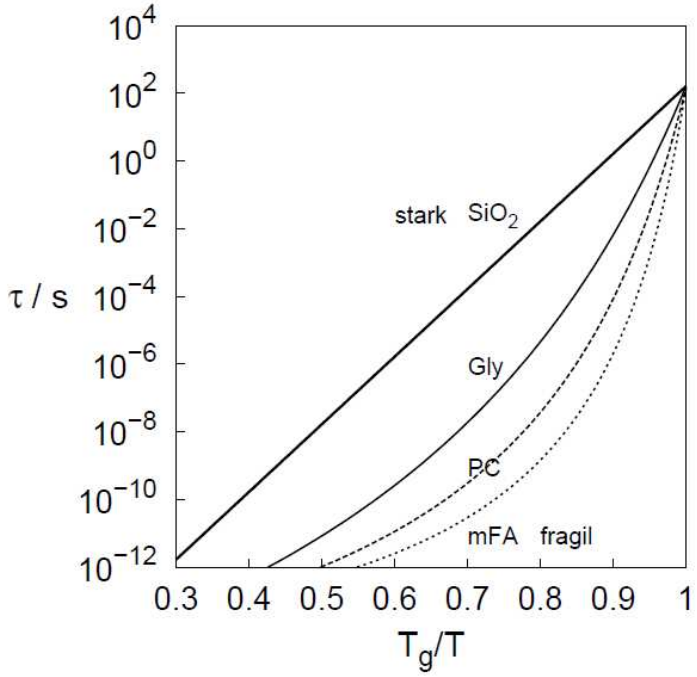


Fig. 5: *The Angell plot shows* $\log \tau_\alpha$ *as a function of* T/T_g*, here for vitreous silica, glycerol and two other glass formers (taken from the thesis of A. Reiser [24]).*

where the factor reflects the sixteen decades between microscopic and macroscopic time scales. I is a better measure of the fragility than m, because it does not contain the trivial temperature dependence of any thermally activated process. If $I = 0$, one has the harmonic case of a temperature-independent energy barrier. Glass formers like vitreous silica with $m = 20$ and $I = 0.25$ are close to this harmonic case and are called strong glass formers, as opposed to to the fragile ones. Glycerol is intermediate with $m = 53$ and $I = 2.3$; the most fragile ones are some polymers with $m = 150$ and $I = 9.4$.
The most popular fitting form for τ_α is the empirical Vogel-Fulcher-Tammann-Hesse relation

$$\log \tau_\alpha = \log \tau_0 + \frac{B}{T - T_0}, \tag{10}$$

which is derived from the Arrhenius equation, eq. (8), by replacing the temperature T in the denominator of the exponent by $T - T_0$. The Vogel-Fulcher temperature T_0 is close to zero in a strong glass and close (but below) T_g in a fragile glass. The Vogel-Fulcher relation predicts a divergence of τ_α at the Vogel-Fulcher temperature. In many glass formers, it equals within the error bars of the experimental determination [26] the thermodynamically defined Kauzmann temperature T_K, suggesting a hidden relation between kinetics and thermodynamics. This will be discussed in more detail in the next section.
Note that the Vogel-Fulcher relation, eq. (10), is known as Williams-Landel-Ferry or WLF-relation in polymers.

2.2 Stretching and dynamical heterogeneity

The second important property of the α-process is its "stretching". The stretching, seen in the time dependence of the α-relaxation, means that there is not only one exponential decay with

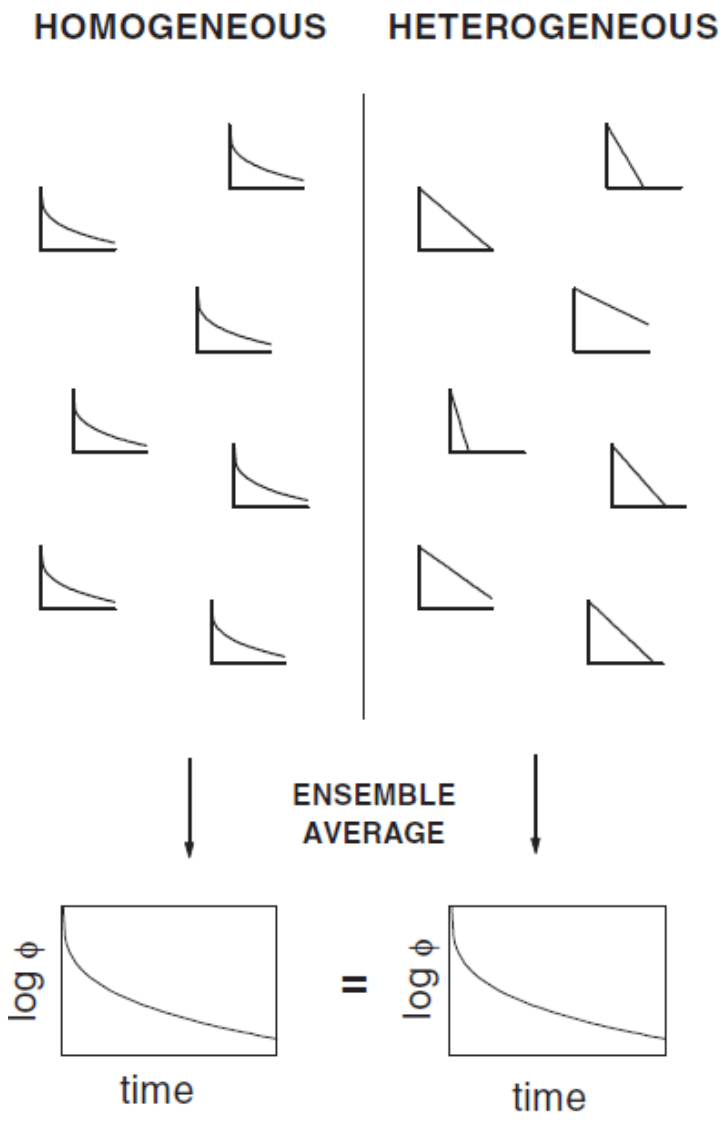


Fig. 6: *The two possible scenarios for the stretching (schematic): On the left dynamical homogeneity, on the right dynamical heterogeneity [28].*

a single relaxation time τ_α, but that one needs a whole distribution of exponential decays with different relaxation times to describe the time dependence of the decay. If one switches on a constant shear strain ϵ_0 in the undercooled sample at the time $t = 0$, the shear stress decay is often well described in terms of a stretched exponential, the so-called Kohlrausch function

$$G(t) = G_{high} \exp(-(t/\tau_\alpha)^\beta). \tag{11}$$

The stretching exponent (the Kohlrausch β) lies between 0.8 and 0.3 (the latter case correspond to a pronounced stretching over three to four decades in time). There is a tendency [27] for a larger stretching for more fragile glass formers, but it is not an exact relation.
If one plots $\log(G(t))$ as a function of time, one sees the stretching in the curvature. $\beta = 1$, the Debye case of a single relaxation time, gives a straight line. The smaller β, the larger the curvature.
The central question with regard to the stretching is whether different regions in the sample have different relaxation times (in this case one talks of dynamical heterogeneity) or whether

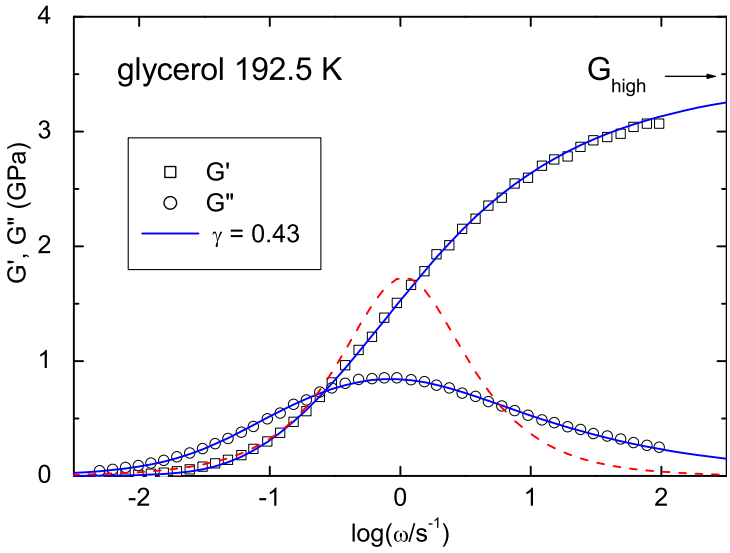


Fig. 7: *Real and imaginary parts $G'(\omega)$ and $G''(\omega)$ of the complex shear modulus in glycerol [29] at 192.5 K as a function of the frequency ω. The continuous line is a fit with a Kohlrausch stretched exponential with $\beta = 0.43$, the dashed line is $G''(\omega)$ for a single exponential decay without any stretching.*

each region of the sample has the same stretched relaxation function (this would be dynamical homogeneity). The two cases are depicted schematically in Fig. 6.

The question can be decided by experiment. The most important technique to do so is NMR, but there are several other possibilities. A review of these techniques and their results has been given by Ranko Richert [28]. It turns out that one has indeed dynamical heterogeneity, at least in the sense that the relaxation functions of different regions in the sample have largely different average relaxation times. Again, we return to the question in the section on new developements.

In most cases, the measurement of the stretching is done in the frequency domain. One can, for example, apply a small-amplitude sinusoidal shear strain $\epsilon_s(t) = \epsilon_0 \cos(\omega t)$ to the sample. Then one needs to exert a sinusoidal shear stress

$$\sigma_s(t) = G'(\omega)\epsilon_0 \cos(\omega t) + G''(\omega)\epsilon_0 \sin(\omega t). \tag{12}$$

$G'(\omega)$ and $G''(\omega)$ are the real and imaginary parts of the complex frequency-dependent shear modulus. $G''(\omega)$, the out-of-phase part of the response, is a measure for the loss at the frequency ω.

Fig. 7 shows the real and imaginary part of the complex shear modulus of glycerol [29] slightly above the glass transition. The comparison with the dashed curve for a single exponential clearly shows a sizable stretching. Glycerol is a Type-A glass former (one distinguishes Type-A = no visible secondary relation peak and Type-B = clearly visible secondary relaxation peak), but one can naturally imagine a small secondary relaxation peak hidden in the extended high-frequency tail. We return to this question in section 4 on new developments.

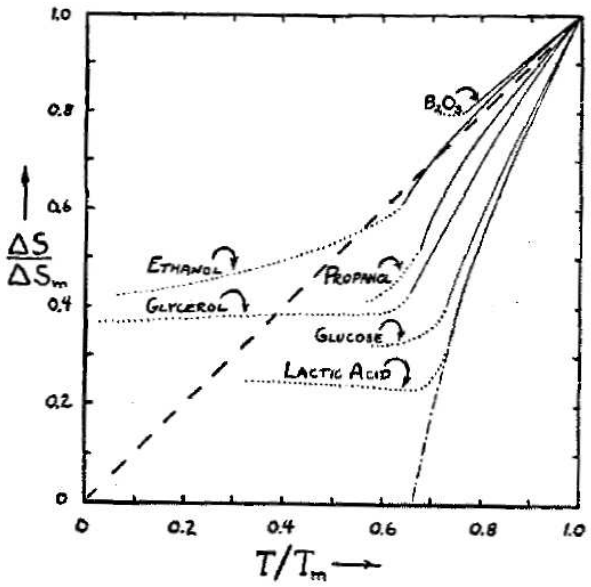


Fig. 8: *Kauzmann's [30] original picture of the loss of the excess entropy $\Delta S = S_{liquid} - S_{crystal}$ on cooling down from the melting temperature in several glass formers.*

3 Thermodynamics of the glass transition

3.1 The Kauzmann catastrophe

Once one is well below the glass transition, the thermodynamics of the glass is rather similar to the one of the crystal. In the example of selenium, the heat capacity c_p of both crystal and glass is close to the classical Dulong-Petit-value of $3k_B$ per atom, due to the three vibrations per atom.

In the undercooled liquid, one has not only the fast energy fluctuations of the vibrations on the picosecond scale, but also the slow energy fluctuations of the structural rearrangements at the relaxation time τ_α. This leads to an additional heat capacity Δc_p. This Δc_p has values between 0.1 and 2 k_B per atom, depending on the specific glass former [6].

Since the heat capacity of the undercooled liquid is higher than the one of the crystal, the undercooled liquid gradually looses the excess entropy $\Delta S = S_{liquid} - S_{crystal}$ over the crystal on cooling. At the melting temperature T_m, $\Delta S = S_m = \Delta H_m/T_m$, where ΔH_m is the latent heat of melting. The melting entropy (which is the physical reason for melting, because it makes the free energy of the liquid more favorable at higher temperatures) is of the order of 0.5 k_B per atom, implying the possibility of about two structural choices per atom in the liquid state. If Δc_p is large, the undercooled liquid rapidly looses this excess entropy on cooling. A linear extrapolation allows to determine the so-called Kauzmann temperature T_K, defined by Kauzmann [30] back in 1948. At the Kauzmann temperature, the entropy of liquid and crystal would be equal, implying essentially only one single possible structural realization of the liquid. It is clear that, at this temperature, the viscosity must diverge, because the liquid becomes unable to jump from one structural realization to another. Fig. 8 shows the Kauzmann extrapolation for some glass formers from his original paper [30]. The figure shows that the glass transition occurs before the undercooled liquid looses all its excess entropy. Thus the "Kauzmann catastrophe" (a liquid with less than one structural configuration) is avoided.

But one can compare the Vogel-Fulcher temperature T_0 (the temperature where the viscosity

would diverge) with the Kauzmann temperature T_K. A comparison to 54 glass formers [26] showed good general agreement, but with four glass formers where T_0 was decidedly smaller than T_K, out of the error bars. Nevertheless, most scientists in the field feel convinced that there is indeed a true physical connection between viscosity and excess entropy. This conviction is further supported by the success of the Adam-Gibbs conjecture [31]

$$\tau_\alpha = \tau_0 \exp\left(\frac{C}{T\Delta S}\right), \tag{13}$$

where C is a constant. The Adam-Gibbs scheme links τ_α and excess entropy via a model of cooperatively rearranging regions, the size of which diverges at the Kauzmann temperature. For a linear dependence of ΔS on temperature, it leads again to the Vogel-Fulcher relation (10) with $T_0 = T_K$.

3.2 The Prigogine-Defay ratio

The question of the nature of the slow structural rearrangements in the undercooled liquid is intimately related to a thermodynamic puzzle, the Prigogine-Defay ratio of the glass transition [32]

$$\Pi = \frac{\Delta c_p \Delta\kappa}{(\Delta\alpha)^2 T_g} = \frac{\overline{\Delta H^2}\ \overline{\Delta V^2}}{(\overline{\Delta H \Delta V})^2}, \tag{14}$$

which relates the increases of the heat capacity per volume unit Δc_p, of the compressibility $\Delta\kappa$ and of the thermal volume expansion $\Delta\alpha$ at the glass temperature T_g to the structural rearrangement enthalpy and volume fluctuations ΔH and ΔV, respectively. If the enthalpy and volume fluctuations are completely correlated, the Prigogine-Defay ratio is one. It has been argued that this is the physically simple case from which one should try to understand the glass transition [33, 34]. We will come back to this proposal in section 4 on new developments.

The Prigogine-Defay ratio is found to be exactly 1 at second order phase transitions (a derivation on the basis of the two Ehrenfest relations for the pressure dependence of the transition temperature is given in Appendix A). A second order phase transition is characterized by continuous first derivations of the enthalpy H, so the volume V and the entropy S are continuous (unlike the first order melting transition, where entropy and volume of liquid and crystal at the melting point are different). At the glass transition, the entropy and the volume are indeed continuous; nevertheless, it is not a second order phase transition, because one of the two phases - the glass - is not in thermal equilibrium. Thus one cannot wonder that real glass formers show Prigogine-Defay ratios much larger than 1 at their glass transition (see Table I). As a rule, stronger glasses have a higher Prigogine-Defay ratio than fragile ones.

subst.	SiO_2	GeO_2	B_2O_3	glycerol	Se	PS	BPA-PC
Π	$>$100	6.85	4.7	3.7	1.85	1.3	1.02
m	20	20	32	53	87	139	132

Table 1: *Measured Prigogine-Defay ratios of seven glass formers at their glass transition [34, 35]. The fragility m is taken from the collection of Böhmer, Ngai, Angell and Plazek [27].*

4 New developments

4.1 Numerical simulations

The most important new development in the field is the application of numerical methods to the glass transition. One can nowadays study the time development of up to hundred thousand interacting particles on the computer. The technique is called "molecular dynamics" and involves the calculation of the motion of the particles in short time steps, short enough to linearize the particle motion. This implies times steps much shorter than one period of the fast vibrations, in real substances of the order of femtoseconds. With a femtosecond, one has to make a million time steps to arrive at a nanosecond. This limits the method at present to relaxation times which are shorter than a microsecond. As a consequence, one cannot study the glass formers in thermal equilibrium at lower temperatures, where τ_α exceeds a microsecond. One just gets slightly below the critical temperature T_c of the mode coupling theory. Thus the first applications of the method have been to test and to verify the mode coupling theory [36, 37]. Later, it has been used to verify a deep and surprising connection [9] between the mode coupling theory and spin glass theory [4], which will be discussed in the next paragraph on new theoretical developments.
In the calculations, one puts the particles in a cubic box with periodic boundary conditions. A particle is free to leave the box (entering the box from the opposite side) and interacts with the particles in the neighboring boxes. The time needed to calculate all particle interactions increases drastically with the range of the potential. Therefore one prefers short-range anharmonic potentials like the Lennard-Jones potential

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], \tag{15}$$

where r is the interparticle distance, ϵ is the depth of the potential well and σ is close to the resulting nearest-neighbor distance (the potential minimum is at $\sigma/2^{1/6}$).
If one just takes the simple Lennard-Jones potential, one finds crystallization to an fcc structure even for the short simulation times. In order to avoid crystallization, it turns out to be necessary to use binary mixtures of particles with different σ. The most popular of these mixtures is the binary Lennard-Jones mixture introduced by Kob and Andersen [36].
With numerical simulations, one can attack the unsolved problems of the glass transition, for instance the question of the Prigogine-Defay ratio. This has been done for the simple Lennard-Jones system [33], calculating not the Prigogine-Defay ratio, but rather the correlation between the enthalpy H and the volume V

$$\Pi^{-1/2} = \frac{\overline{HV}}{\sqrt{\overline{H^2}\;\overline{V^2}}}, \tag{16}$$

a correlation which involves not only the slow structural degrees of freedom, but the fast vibrations as well.
It is easier to do this in a constant volume calculation at constant temperature (the NVT ensemble), looking at the correlation between pressure fluctuations and energy fluctuations. In this calculation, one has to subtract the ideal gas term, because the pressure p is given by

$$pV = Nk_BT + W, \tag{17}$$

where only the second term, the virial W, is due to the interatomic potential. As one can see in Fig. 9, the virial is indeed highly correlated with the potential energy, with a correlation factor

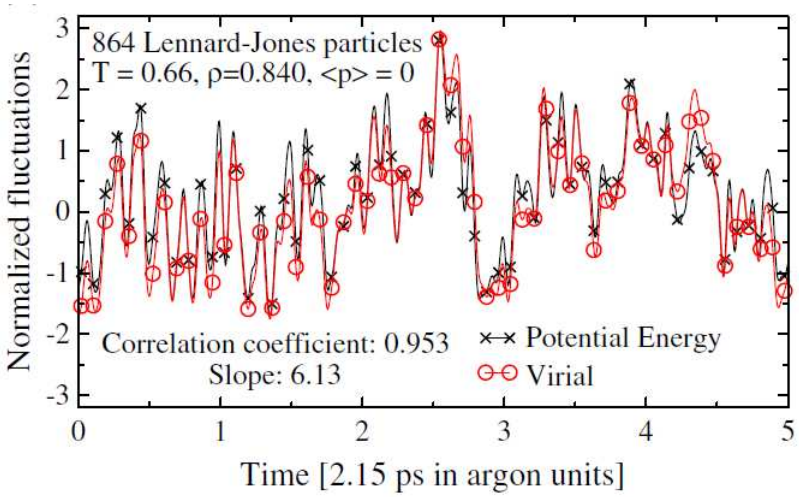


Fig. 9: *Instantaneous normalized equilibrium fluctuations of virial W and potential energy U in the standard single-component LennardJones (LJ) liquid at constant volume (NVT simulation), showing that W(t) and U(t) correlate strongly [33].*

0.953 corresponding to a Prigogine-Defay ratio of 1.1. The same high correlation is found in all simulations of Van der Waals and metallic liquids, but not in hydrogen-bonded liquids like water and glycerol.

On the basis of these results, it has been argued that a Prigogine-Defay ratio of 1 is the physically simple case from which one should try to understand the glass transition [33, 34]. But this is opposed to the arguments in the introduction of the present paper, which indicate that the physically simple case is a harmonic glass former. In a harmonic glass former, the thermal expansion should be zero. Therefore the linear correlation $\overline{\Delta H \Delta V}$ should be zero; the Prigogine-Defay ratio should be infinity (as one indeed finds in silica [38] with $\Pi > 100$).

What brings the Prigogine-Defay ratio down, is the anharmonicity of the potential, which also causes the thermal expansion. With anharmonicity, even a pure shear energy fluctuation is accompanied by a volume fluctuation. If these additional volume fluctuations dominate, the Prigogine-Defay ratio goes to one. In numerical work [36, 33], one usually prefers anharmonic potentials, which vanish at a relative close distance from the atoms. Thus one tends to conclude that the deviation of the Prigogine-Defay ratio from one is an indication for a complicated atomic potential [33, 34]. But it is rather an indication for a harmonic atomic potential.

This point of view is supported by Fig. 10, taken from a very recent paper [40]. As a measure for the anharmonicity at the glass temperature T_g, one can use the product $\alpha_l T_g$, where α_l is the volume expansion coefficient of the liquid. Fig. 10 shows the correlation coefficient $\Pi^{-1/2}$ as a function of this product. Though again the experimental data [34, 35] scatter strongly, one finds clear evidence for a linear increase at low anharmonicity. The simulation value for the zero-pressure binary Kob-Andersen mixture [33] corresponds to rather high anharmonicity.

Fig. 10 demonstrates the limitations of numerical simulations: They do not get very low in temperature and they tempt scientists to overestimate the importance of anharmonic potentials. In spite of these limitations, numerical simulations were at the very bottom of the astonishing progress in the understanding of the glass transition within the last decade, which will be described in the next paragraph.

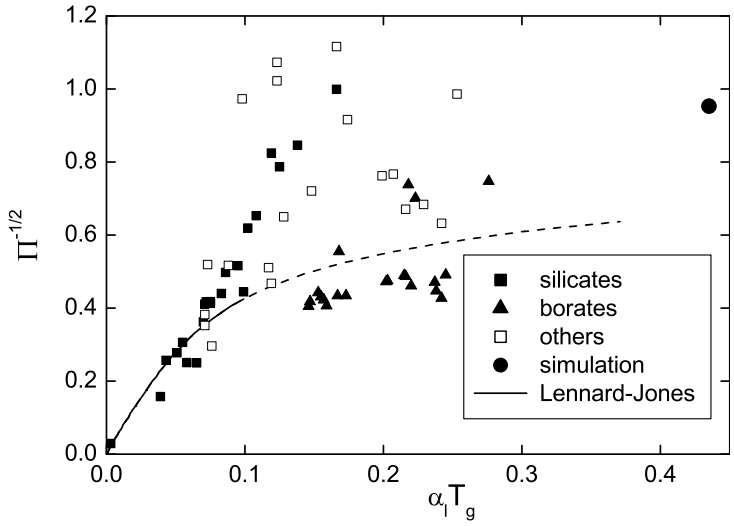


Fig. 10: *The correlation coefficient* $\Pi^{-1/2}$ *(*Π *Prigogine-Defay ratio) as a function of the anharmonicity* $\alpha_l T_g$ *(*α_l *volume expansion coefficient of the liquid,* T_g *glass temperature) for 55 glass formers [34, 35, 39] and the binary Kob-Andersen mixture [33]. The line is a numerical calculation for the Lennard-Jones potential in ref. [40], from which the figure was taken.*

4.2 Theoretical developments

The most important theoretical development of the last decade was the discovery of a deep analogy between mode coupling theory [3] and spin glass theory [4]. There is an excellent recent review of the topic [9], which demonstrates the intimate connection between theory and numerical simulation in this development.

The merit of the mode coupling theory [3] is not so much its prediction of a divergence of the viscosity at T_c - the viscosity does not really diverge there, because thermally activated processes take over - but its prediction of a slowing down of the flow relaxation time τ_α from the vibrational picosecond time scale as one lowers the temperature towards T_c. According to the mode coupling theory, this separation of time scales occurs *without* any activated process. As one approaches T_c, the structural correlation functions decay in two steps, the first in the picosecond regime and the second at τ_α.

However, since the mathematical apparatus of the mode coupling theory is rather complicated, it is not easy to assess a clear physical significance to the separation of time scales. It was said that τ_α corresponded to the "breaking of the cage of neighboring particles", but this is not much more than a figure of speech.

A much deeper explanation became possible when it was realized that the dynamic mode coupling equations are the same as those for the mean field p-spin model [41], a spin-glass model with the Hamiltonian

$$H = -\Sigma_{i,j,k=1}^{N} J_{ijk}\sigma_i\sigma_j\sigma_k \tag{18}$$

for $p = 3$, where J_{ijk} are frozen random variables and σ_i, σ_j and σ_k are spin variables. Since each spin interacts which each pair of other spins irrespective of their distance, the model contains no space information.

The advantage of the p-spin model is that it allows to calculate all important properties. Thus one can analyze what happens as one approaches T_c. One finds [42] that the separation of times scales is due to the properties of the saddle points of the system, which have the tendency to

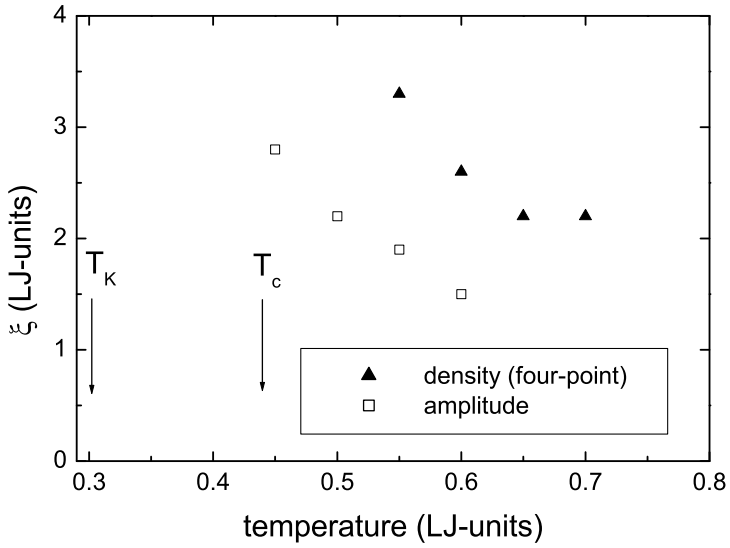


Fig. 11: *Simulation results for a diverging length scale in the binary Kob-Andersen Lennard-Jones mixture from the 4-point correlation function [45] and from the amplitude correlation in shear-induced inherent structure changes [46].*

become isolated true free energy minima toward T_c. This makes the way between two given equilibrium configuration of the system longer and longer as one approaches T_c. Thus the separation of time scales predicted by the mode coupling theory can be identified quantitatively as a phase space property [9].

Scarcely less important, the replica technique (sometimes also called "cloning theory") developed in the context of spin glass theory [4] allows to do calculations for the undercooled liquid below T_c. Thus one calculate the heat capacity [43] for a given interatomic potential between T_c and the Kauzmann temperature T_k. Similarly, one finds the temperature dependence of the plateau shear modulus G_{high} between these two temperatures [15], concluding that it decreases to zero as one reaches T_c.

Again in the same context, the spin glass ideas have been extended to structural glasses in an intense search for a dynamic length scale [44, 45, 46] which is supposed to diverge at the Kauzmann temperature T_k. There is increasing numerical evidence [45, 46] (see Fig. 11) for such a diverging length scale in the binary Kob-Andersen Lennard-Jones mixture, supporting the Adam-Gibbs concept [31] of cooperatively rearranging regions. The motivation of the work is to clarify the deep connection between thermodynamics and dynamics reflected in the validity of the Adam-Gibbs equation (13).

A theoretical scheme which indeed leads to a validity of the Adam-Gibbs equation is the mosaic theory or random-first-order theory [47] (RFOT), motivated by the parallel between mode coupling theory and spin glass models. According to RFOT, the diameter of a cooperatively rearranging region results from the thermodynamic equilibrium between the entropy inside the region and the surface tension at the interface to neighboring regions. One has to postulate a surface tension coefficient which depends on the diameter to get the Adam-Gibbs equation, but one can find a theoretical justification for this [47].

The theoretical evolution described so far seems fairly coherent and convincing, but there are other and independent theoretical attempts to understand the glass transition. Some of them are described at the end of Cavagna's excellent review [9]. Others are motivated by new experi-

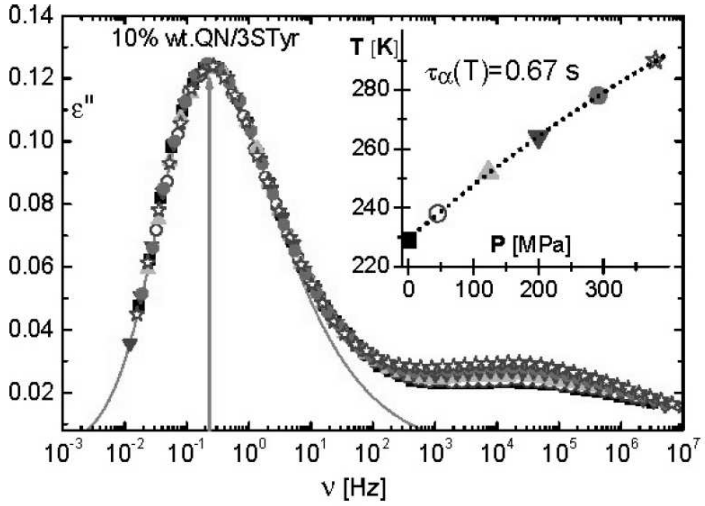


Fig. 12: *Dielectric spectra of 10 mol%-quinaldine in tristyrene at different pressures and adapted temperatures [48]. The inset shows the curve in the temperature-pressure diagram which must be followed to keep the primary peak frequency at the same place.*

mental findings and will be described in the next paragraph.

4.3 New experimental findings

Though a bit overshadowed by the success of the numerical simulations described above, the progress in experimental techniques for the study of the glass transition is by no means contemptible. The main progress has been made by the development of broadband dielectric spectroscopy [49, 50]. It is nowadays possible to measure the dielectric spectrum from 10^{-6} to 10^{12} Hz (though the GHz region is still not easy to measure). Even more important, one can measure the dielectric spectrum under high pressure. If one does this for glycerol, which has no clearly visible secondary relaxation at ambient pressure, one does indeed see the development of a clear secondary peak [51].

Fig. 12 shows a measurement of a binary mixture, 10-mol% of quinaldine in tristyrene [48]. One sees the large and broad α- peak at low frequency and a smaller secondary peak at higher frequency. This is a clever pressure experiment, varying temperature and pressure simultaneously in such a way that the primary peak stays at the same frequency. One finds the striking result that the secondary peak also stays at the same place, indicating a connection between α and β-process.

However, the picture can also be quite different [52]. In some systems, where the molecules are not rigid and the secondary relaxation is due to to a molecular configuration change, the primary peak is much more sensitive to pressure than the secondary peak. So it is necessary to distinguish these two cases. If one has a molecular configuration change, the two inherent states which contribute to the secondary relaxation will be separated by a practically temperature- and pressure-independent barrier. In the other case, one has a barrier between two different molecular arrangements, which will vary with temperature and pressure according to the variation of the elastic constants, in particular the high frequency shear modulus G_{high} (in fact, the elastic models [10] explain the fragility by a proportionality of the strongly temperature-dependent G_{high} to the flow barrier).

Taking the pressure dependence as an indicator whether one deals with an intramolecular or

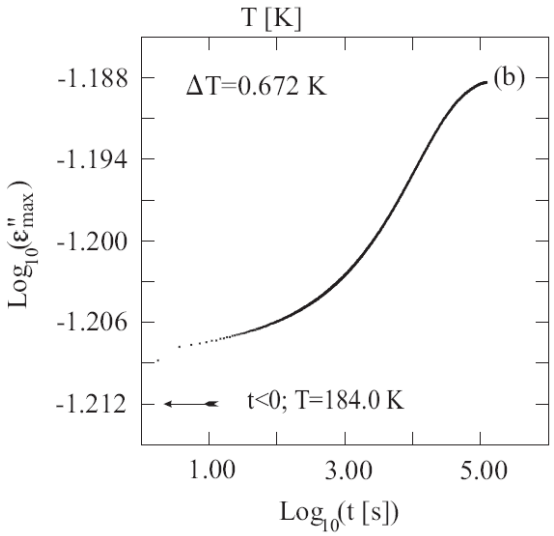


Fig. 13: *Beta loss peak maximum in tripropylene glycol, monitored after a temperature increase of 0.672 K starting from equilibrium at 184.0 K. After 6 s temperature is stable within 1 mK of the final temperature. At this time, the energy landscape has not yet changed, but the peak maximum nevertheless shows already a clearly measurable increase compared to the value before heating, denoted by the arrow [55].*

intermolecular relaxation, one finds [48, 52, 53] that the relaxation time τ_β of the latter ones always follows the coupling model relation [53]

$$\tau_\beta = \tau_\alpha^{1-n} t_c^n, \tag{19}$$

where $1-n$ is the Kohlrausch exponent of the primary process and t_c is a short time of the order of 2 ps. According to the coupling model, the α-process reflects the result of the interaction of the "primitive" relaxations in the β-process. This is a picture of the glass transition which is completely different from the mode-coupling and spin-glass concept described in the preceding two paragraphs, but there are many dielectric experiments supporting it [11].

In fact, a connection between α- and β-process has been demonstrated directly in another key experiment [54], done by NMR on sorbitol. This experiment is one of the many measurements demonstrating dynamical heterogeneity [28], i.e. different relaxation times at different places in the sample. The NMR technique can be used to separate subensembles of the sample which belong to a specific relaxation time. The experiment showed that a subensemble which has a slow α-relaxation time also tends to have a slow β-relaxation time. As the authors [54] point out, there is at present only a single theoretical explanation for such a behavior, namely Kia Ngai's coupling model [53].

The last experiment reported here is a unique aging experiment [55] close to the glass temperature, where the relaxation time τ_α is of the order of a day. It is done within a sample holder for dielectric measurements which is able to change the temperature to a new equilibrium within a few seconds. This allows to see what happens with the fast secondary relaxation during the aging process.

Fig. 13 shows the development of the amplitude of the secondary process in tripropylene glycol at 184 K, where the primary process has a relaxation time of a day. After a quick heating to the temperature 184.672 K, one follows the gradual approach to equilibrium at long times.

But the surprising thing is: One already sees a small, but decidedly measurable increase of the amplitude at short times, after the equilibration time of the temperature. This means an increase within an unchanged energy landscape, and this in turn has only one possible explanation: The inherent states, between which the system jumps back and forth in the secondary relaxation, must be strongly asymmetric. One can calculate the average asymmetry from the increase. For the one in Fig. 12, one finds an average asymmetry of 3.8 $k_B T_g$, a factor four or more larger than one would expect.

This is not a specific property of tripropylene glycol, because the authors measured it in several systems, including a toluene-pyridine mixture of rigid molecules. It seems to be a general phenomenon, irrespective of whether the secondary peak is a molecular rearrangement of rigid molecules or an intra-molecular configuration change.

An explanation of this asymmetry was proposed by the author [56] in terms of the Eshelby backstress [2]. If the system searches for possible neighboring inherent states, it will most probably find a mayority of elastically distorted states, which do not fit well into the present strain of the surrounding matrix. The concept allows to predict an asymmetry of 4 $k_B T$ without any adaptable parameter, close to the observed value in tripropylene glycol.

If this hypothesis is generally true, the α-relaxation should show the same asymmetry. Some of the implications of such a general asymmetry have been worked out in a series of papers [40, 56, 57].

5 Summary

The glass transition is at present a very active field of research. The main reason for this is the very fruitful cooperation of theory and numerical simulation described in Section 4 of this paper [9]. It encourages the hope for a closed quantitative and convincing description of the glass transition within the next decade. In this description, there is a separation of time scales on cooling the undercooled liquid, which occurs when the valleys in the free energy landscape are no longer connected. The onset of this separation of time scales is well described by the mode coupling theory [3]. As one approaches the critical temperature of the mode coupling theory from above, thermally activated hopping between the minima of the free energy landscape (at this temperature already practically the potential energy landscape [5, 7]) sets in. In this temperature regime, the glass formers approach a true thermodynamic second order transition at the Kauzmann temperature. The more fragile the glass former is, the closer it gets to the Kauzmann temperature. In this regime, one can calculate physical properties quantitatively if one knows the interatomic potentials, using the replica or cloning theory [15, 43].

On the other hand, there are important experimental facts which remain unexplained and are not (or not yet) integrated into this appealing picture: the role of the anharmonicity [40], the relations between primary and secondary relaxation processes which seem to be well described by the coupling model [11] and the asymmetry of the secondary relaxation [55], which can be argued [56] to be a general property due to the Eshelby backstress [2].

Appendices

A The Prigogine-Defay ratio at a second order phase transition

The differential of the Gibbs free energy G is given by

$$dG = -SdT - Vdp, \tag{20}$$

where S is the entropy, T the temperature, V the volume and p the pressure of the system. From this equation follows the Maxwell relation

$$\frac{\partial^2 G}{\partial T \partial p} = -\left(\frac{\partial S}{\partial p}\right)_T = \left(\frac{\partial V}{\partial T}\right)_p = -V\alpha, \tag{21}$$

where α is the thermal expansion, because the order of the two diffentiations of G is irrelevant. Now consider a second order phase transition between phases A and B, in which entropy and volume are continuous. The two phases join at the transition temperature $T_c(p)$, which depends on the pressure and thus yields a transition line in a p, T-diagram. Along this line, one has two continuity relations, the first from the entropy

$$\left(\frac{\partial S_A}{\partial T}\right)_p dT + \left(\frac{\partial S_A}{\partial p}\right)_T dp = \left(\frac{\partial S_B}{\partial T}\right)_p dT + \left(\frac{\partial S_B}{\partial p}\right)_T dp \tag{22}$$

and the second from the volume

$$\left(\frac{\partial V_A}{\partial T}\right)_p dT + \left(\frac{\partial V_A}{\partial p}\right)_T dp = \left(\frac{\partial V_B}{\partial T}\right)_p dT + \left(\frac{\partial V_B}{\partial p}\right)_T dp. \tag{23}$$

The derivatives of S with respect to p can be replaced by the Maxwell relation (21). $(1/V)\partial S/\partial T$ at constant pressure is the heat capacity c_p per volume unit, $-(1/V)\partial V/\partial p$ at constant temperature is the isothermal compressibility κ_T. With these definitions, and setting Δc_p, $\Delta\alpha$ and $\Delta\kappa_T$ for the differences of phase A and phase B in heat capacity, thermal expansion and compressibility, respectively, one obtains the two Ehrenfest relations

$$\frac{\partial T_c}{\partial p} = \frac{T\Delta\alpha}{\Delta c_p} \tag{24}$$

and

$$\frac{\partial T_c}{\partial p} = \frac{\Delta\kappa_T}{\Delta\alpha}. \tag{25}$$

Equating these two, one finds that the Prigogine-Defay ratio of eq. (14) is indeed 1. However, as pointed out above, this equality requires thermal equilibrium in both phases A and B. For the glass transition, one can argue [9] that one reaches this requirement for $T_g = T_K$, which in turn helps to understand why the Prigogine-Defay ratio is indeed close to 1 for very fragile glass formers (see Table I).

References

[1] C. Kittel, *Introduction to Solid State Physics*, 3rd ed., (Wiley, New York 1966), p. 183. This reference does not say anything about the glass transition, but explains the anharmonicity mechanism of the thermal expansion.

[2] J. D. Eshelby, Proc. Roy. Soc. **A241**, 376 (1957). Again, this reference says nothing about the glass transition, but gives the elasticity theory description of a structural rearrangement within an isotropic elastic solid.

[3] W. Götze and S. Sjögren, Rep. Prog. Phys. **55**, 241 (1992). This reference explains the mode coupling theory.

[4] M. Mezard, G. Parisi and M.A. Virasoro, *Spin-Glass theory and beyond* (World Scientific, Singapore, 1987)

[5] M. Goldstein, J. Chem. Phys. **51**, 3728 (1969). The energy landscape concept.

[6] C. A. Angell, W. Sichina, Ann. New York Acad. Sci. **279**, 53 (1976). Experimental results at the glass transition.

[7] P. G. Debenedetti and F. H. Stillinger, Nature **410**, 259 (2001). The concept of inherent structures.

[8] F. Leonforte, R. Boissiere, A. Tanguy, J. P. Wittmer and J.-L. Barrat, Phys. Rev. B **72**, 224206 (2005). Nonaffine atomic motion after an affine shear deformation of an amorphous solid.

[9] A. Cavagna, Phys. Rep. **476**, 51 (2009). Review of (a) the connection between mode coupling theory and spin glass theory (b) thermodynamic calculations below the critical temperature of the mode coupling theory with the help of the spin glass theory.

[10] J. C. Dyre, Rev. Mod. Phys. **78**, 953 (2006). A review of the elastic models.

[11] K. L. Ngai, *Relaxation and Diffusion in Complex Systems* (Springer, Berlin 2011). Describes the experimental evidence for the coupling model.

[12] U. Buchenau, J. Non-Cryst. Solids **357**, 274 (2011). Key experiments in highly viscous liquids.

[13] S. V. Nemilov, *Thermodynamic and Kinetic Aspects of the Vitreous State*, (CRC Press, Ann Arbor 1995)

[14] E. Donth, *The Glass Transition*, (Springer, Berlin 2001)

[15] H. Yoshino and M. Mezard, Phys. Rev. Lett. **105**, 015504 (2010)

[16] U. Buchenau and R. Zorn, Europhys. Lett. **18**, 523 (1992)

[17] J. Wuttke, W. Petry, G. Coddens and F. Fujara, Phys. Rev. E **52**, 4026 (1995)

[18] F. Scarponi, L. Comez, D. Fioretto and L. Palmieri, Phys. Rev. B **70**, 054203 (2004)

[19] U. Buchenau, Yu. M. Galperin, V. L. Gurevich, D. A. Parshin, M. A. Ramos and H. R. Schober, Phys. Rev. B **46**, 2798 (1992)

[20] H. G. E. Hentschel,S. Karmakar, E. Lerner and I. Procaccia, Phys. Rev. E **83**, 061101 (2011)

[21] W. Schirmacher, G. Ruocco and T. Scopigno, Phys. Rev. Lett.**98**, 025501 (2007)

[22] W. A. Phillips, U. Buchenau, N. Nücker, A. J. Dianoux and W. Petry, Phys. Rev. Lett. **63**, 2381 (1989)

[23] A. Wischnewski, U. Buchenau, A. J. Dianoux, W. A. Kamitakahara, and J. L. Zarestky, Phys. Rev. B **57**, 2663 (1998)

[24] A. Reiser, Ph. D. thesis, Univ. Heidelberg 2005

[25] C. A. Angell, Science **267**, 1924 (1995)

[26] L.-M. Wang, C. A. Angell and R. Richert, J. Chem. Phys. **125**, 074505 (2006)

[27] R. Böhmer, K. L. Ngai, C. A. Angell and D. J. Plazek, J. Chem. Phys. **99**, 4201 (1993)

[28] R. Richert, J. Phys.: Condens. Matter **14**, R703 (2002)

[29] K. Schröter and E. Donth, J. Chem. Phys. **113**, 9101 (2000)

[30] A. W. Kauzmann, Chem. Rev. **43**, 219 (1948)

[31] G. Adam and J. H. Gibbs, J. Chem. Phys. **43**, 139 (1958)

[32] J. Jäckle, J. Chem. Phys. **79**, 4463 (1983)

[33] U. R. Pedersen, N. Gnan, N. P. Bailey, T. B. Schroeder and J. C. Dyre, J. Non-Cryst. Solids **357**, 320 (2011)

[34] D. Gundermann, U. R. Pedersen, T. Hecksher, N. P. Bailey, B. Jakobsen, T. Christensen, N. B. Olsen, T. B. Schroeder, D. Fragiadakis, R. Casalini, C. M. Roland, J. C. Dyre and K. Niss, Nature Physics **7**, 816 (2011)

[35] D. B. Dingwell, R. Knoche and S. L. Webb, Phys. Chem. Minerals **19**, 445 (1993)

[36] W. Kob and H.C. Andersen, Phys. Rev. E **52**, 4134 (1995)

[37] H. R. Schober, Phys. Rev. Lett. **88**, 145901 (2002)

[38] D. M. Krol, K. B. Lyons, S. A. Brawer and C. R. Kurkjian, Phys. Rev. B **33**, 4196 (1986)

[39] S. V. Nemilov, V. N. Bogdanov, A. M. Nikonov, S. N. Smerdin, A. I. Nedbai and B. F. Borisov, Fiz. i Khim. Stekla **13**, 801 (1987) (Sov. J. Glass Phys. Chem. **13**, 413 (1987))

[40] U. Buchenau, arXiv:1110.5066 (cond-mat:dis-nn]

[41] J.-P. Bouchaud, L. Cugliandolo, J. Kurchan and M. Mezard, Physica A **226**, 243 (1996)

[42] A. Cavagna, I. Giardina and G. Parisi, Phys. Rev. B **57**, 11251 (1998)

[43] M. Mezard and G. Parisi, J. Phys. A: Math. Gen. **29**, 6515 (1996)

[44] L. Berthier, G. Biroli, J. P. Bouchaud, W. Kob, K. Miyazaki and D. R. Reichman, J. Chem. Phys. **126**, 184503 (2007)

[45] S. Karmakar, C. Dasgupta and S. Sastry, Phys. Rev. Lett. **105**, 015701 (2010)

[46] M. Mosayebi, E. Del Gado, P. Ilg and H. C. Öttinger, Phys. Rev. Lett. **104**, 205704 (2010)

[47] T.R. Kirkpatrick, D. Thirumalai and P.G. Wolynes, Phys. Rev. A **40**, 1045 (1989)

[48] K. Kessairi, S. Capaccioli, D. Prevosto, M. Lucchesi, S. Sharifi, and P. A. Rolla, J. Phys. Chem. B **112**, 4470 (2008)

[49] U. Schneider, P. Lunkenheimer, R. Brand and A. Loidl, J. Non-Cryst. Solids **235-237**, 173 (1998)

[50] *Broadband dielectric Spectroscopy*, ed. F. Kremer and A. Schönhals (Springer, Berlin, 2002)

[51] A. A. Pronin, M. V. Kondrin, A. G. Lyapin, V. V. Brazhkin, A. A. Volkov, P. Lunkenheimer and A. Loidl, Phys. Rev. E **81**, 041503 (2010)

[52] K. L. Ngai and M. Paluch, J. Chem. Phys. **120**, 857 (2004)

[53] K. L. Ngai, Comments Solid State Phys. **9**, 121 (1979); J. Phys.: Condens. Matter **15**, 1107 (2003)

[54] R. Böhmer, G. Diezemann, B. Geil, G. Hinze, A. Nowaczyk and M. Winterlich, Phys. Rev. Lett. **97**, 135701 (2006)

[55] J. C. Dyre and N. B. Olsen, Phys. Rev. Lett. **91**, 155703 (2003)

[56] U. Buchenau, J. Chem. Phys. **131**, 074501 (2009)

[57] U. Buchenau, J. Chem. Phys. **134**, 224501 (2011)

E 10 Protein Crystallography

T. E. Schrader
Jülich Centre for Neutron Science
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Apart from water proteins are the most abundant molecules in living cells. They are constantly synthesized in the cell by the well known transcription and translation mechanism where first the DNA is read out to produce a messenger RNA which encodes the proteins and in the subsequent translation process the protein is synthesized by the ribosomes which itself is a protein/RNA-complex. Proteins fulfil numerous functions in the cell, for example enzymatic catalysis which enhances the speed with which molecules like fatty acids are synthesized or transport and storage of important molecules like oxygen or they are involved in immunology, just to name a few [1]. In order to perform these functions proteins adopt a unique three dimensional structure with a carefully controlled mixture of flexibility and stiffness. For an understanding of their function knowledge of this three dimensional structure is a prerequisite. Ideally one would like to produce a movie where one can follow the functioning protein in action in slow motion with atomic resolution. In practice, there are techniques available which have a sufficient time resolution (in the fs-regime) but do only provide very limited structural information like time resolved infrared spectroscopy. On the other hand there are methods which provide full atomic resolution but with essentially no time resolution. With those methods one often stops the functioning process of a protein under investigation in an intermediate state by trapping methods using inhibitor molecules which stop the catalytic process of the protein leaving it trapped it in a certain intermediate state.
This article will focus on the latter static techniques among which x-ray protein crystallography is the most widely used one. It will also introduce the method of neutron protein crystallography since there are some similarities but also some differences to x-rays as probes. Finally an example case study is discussed where both techniques give complimentary information. But at first a short introduction into the basic structural properties of proteins is given.

2 Some Basics about Proteins

2.1 Amino acids as the building blocks of proteins

Proteins consist of a chain of amino acids. In that sense they are biopolymers and since they in general have charged side chains they can also be called polyelectrolytes. The first information about a protein is therefore the number of amino acid residues it contains. This number can span quite a wide range between 10 and 25 000. In a historical nomenclature often the term “polypeptide” is used for a small protein containing between 10 and 100 amino acids. Amino acid chains with a smaller number of amino acids than 10 are often named oligopeptides. Typically a protein contains 100 amino acids. The average molecular weight per amino acid is around 100 g/mol. This provides a possibility to calculate an estimate for the molecular weight of a protein from the number of its residues. Despite there are many different amino acids (or to be more exact: 2-amino carboxylic acids) present in living organisms not all of them are used to build up proteins. The proteinogenic amino acids are shown in Figure 1. Covalently attached to the central C-atom, the so called C_α atom are an amino functional group, a carboxylic group, the side chain atoms and finally one hydrogen atom. Since the C_α atom has a sp^3 hybridisation sterically all four constituents point into the corners of a tedrahedron. Since (apart from the amino acid glycine) this C_α atom has four

different constituents it forms a chiral centre in the Fischer sense. So two different arrangements of these four constituents are possible which lead to the L- and D-enantiomers of the corresponding amino acid, according to Fischer's convention. But in nature only the L-enantiomers are found as building blocks of proteins.

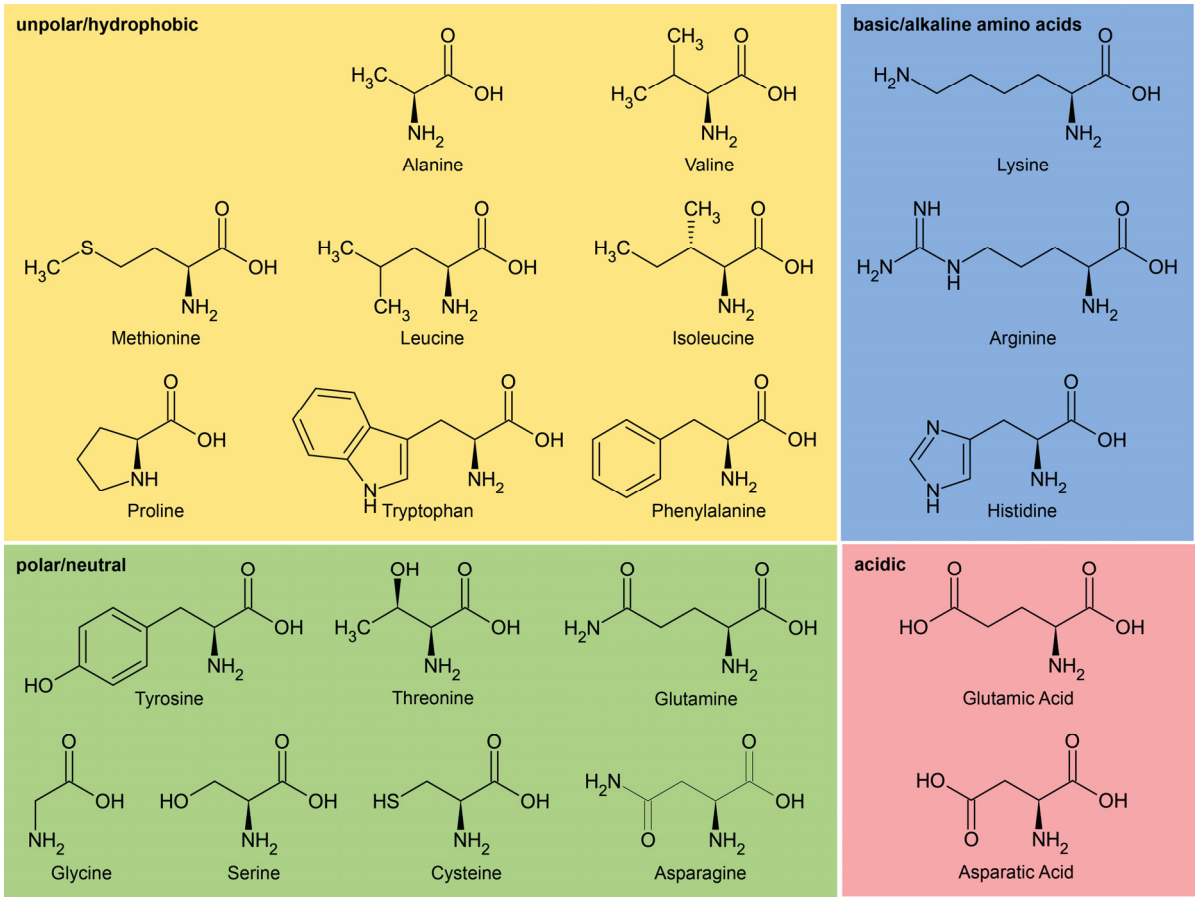


Fig. 1: *A compilation of all 20 amino acids found in natural proteins. The N-terminal amino group is here shown in its neutral charge state pointing to the bottom of the page. Covalently attached to it is the Cα atom which carries the corresponding side chain group (Adapted from http://upload.wikimedia.org/wikipedia/commons/7/7d/Overview_proteinogenic_amino_acids-DE.sv).*

The ribosomes in a living cell synthesize the proteins according to the code read from the messenger RNA. This process is called translation[1]. Hereby in a step by step fashion one amino acid after the other is linked via a peptide bond between the carboxylic group of the existing amino acid chain and the amino functional group of the newly added amino acid. A water molecule is released per peptide bond formed (see Fig. 2). Hence, the peptide bond formation is a poly-condensation. The inverse process requires adding a water molecule and is named hydrolysis. Formation of a peptide bond requires free energy, so the inverse process is

[1] The term transcription is used to describe the production of the messenger RNA by reading the DNA code.

exothermic but happens on a very long time scale between 10 to 1000 years without enzymatic catalysis.

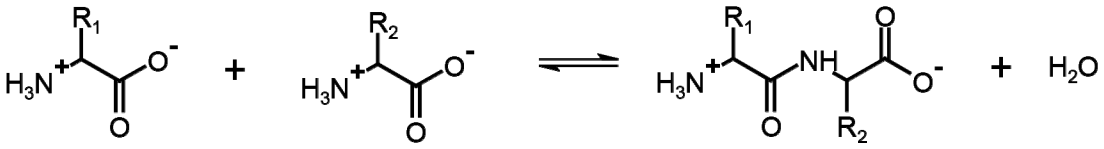


Fig. 2: *A peptide bond forms between two amino acids. As a result a water molecule is released. The inverse process is called hydrolysis.*

The sequence of amino acids is the **primary structure** of the protein. It can be displayed as a line of text with a three letter code representing one amino acid. It is a common convention that the line of text starts at the N-terminal end of the amino acid chain i. e. with the amino acid with side chain R_1 in the example given in Fig. 2 on the right.

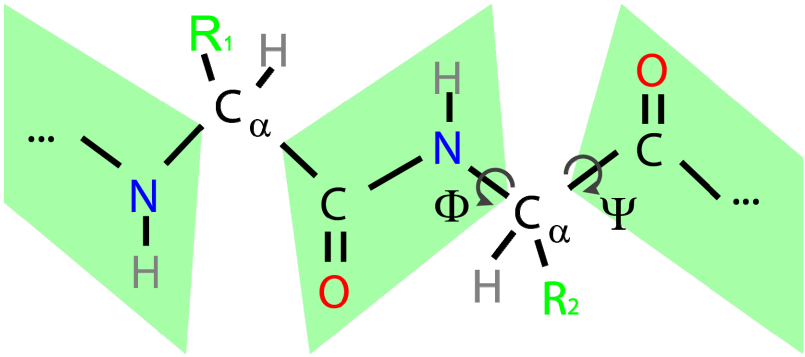


Fig. 3: *Due to a partial double bond character of the C-N peptide bond the rotation around it is hindered by a steep potential. This causes the four atoms OCNH to form a planar structure (marked in green). The only remaining degrees of freedom per amino acid along the backbone are the two dihedral angles Φ and Ψ.*

Another specialty of the peptide bond is its partial double bond character of the chemical bond between the carbon and nitrogen atom. The lone electron pair at the nitrogen atom is delocalized and has some existence probability between the atoms forming the peptide bond. This causes the bond length to shrink below the value of a single bond. As a consequence, the rotation around the CN-bond is hindered and the four atoms OCNH form a planar geometry denoted by a green polygon in Fig. 3[2]. This is why one amino acid only contributes two degrees of freedom to the amino acid backbone which are denoted by the dihedral angles Φ and Ψ, for their definition see Fig. 3. Another property of this planar set of atoms is their potential to from hydrogen bonds. Hereby, the oxygen atom is partly negatively charged and

[2] Only in rare cases one finds a distorted planar geometry. The dihedral angle ω for the rotation around the CN-bond is in this case not equal to +180°. The potential for rotation around the CN-bond has a second local minimum at ω=0°, which corresponds to the peptide bond in its cis configuration. This is frequently found in conjunction with the amino acid proline.

is a hydrogen bond acceptor. The hydrogen atom covalently linked to the backbone nitrogen atom carries a positive partial charge and is therefore a hydrogen bond donor. It is this hydrogen atom which can be easily replaced by a deuterium atom when the protein is dissolved in heavy water (D_2O) for a certain time. This can be seen as a proof for the hydrogen donor capabilities of this hydrogen atom.

2.2 The three dimensional structure of proteins

After leaving the exit tunnel of the ribosome the polypeptide chain folds into a unique three dimensional structure. This process is sometimes assisted by chaperones, which provide a special electrostatic environment, which helps the proteins to fold correctly. Since the backbone of all proteins is the same (i. e. the covalently linked atoms N- C_α -C-N- C_α -C and so forth) the side chains determine this unique three dimensional structure. This structure is stabilized by four different interactions. First of all there is the possibility of establishing **hydrogen bonds** between two parts of the backbone, but also between side chains or between a side chain and a part of the backbone. Another stabilizing mechanism is a formed **salt bridge** between a negatively and a positively charged side chain, e. g. aspartic acid and lysine. The third interaction is the formation of a **hydrophobic cluster** or core. Hereby the surrounding water plays a major role which makes it a mostly entropic effect. It is more favorable for the water molecules to form hydrogen bonds with each other than to stick between some hydrophobic side chains. This is why those side chains tend to be packed together in the folding process resulting in van der Waals interactions among them. The fourth stabilizing moment of a three dimensional fold of a protein is a **disulfide bridge** between two cysteine residues. Often this is used to link two different amino acid chains to form one protein.

When the primary structure only gives the linear sequence of amino acids, the **secondary structure** of a protein denotes all arrangements of the protein backbone stabilized by a regular hydrogen bonding pattern. These hydrogen bonds are solely between different parts of the backbone. There are several structural motifs of that kind which occur frequently in proteins. Some of these motifs have been given a name e .g α-Helix or parallel β-sheet (Fig. 4).

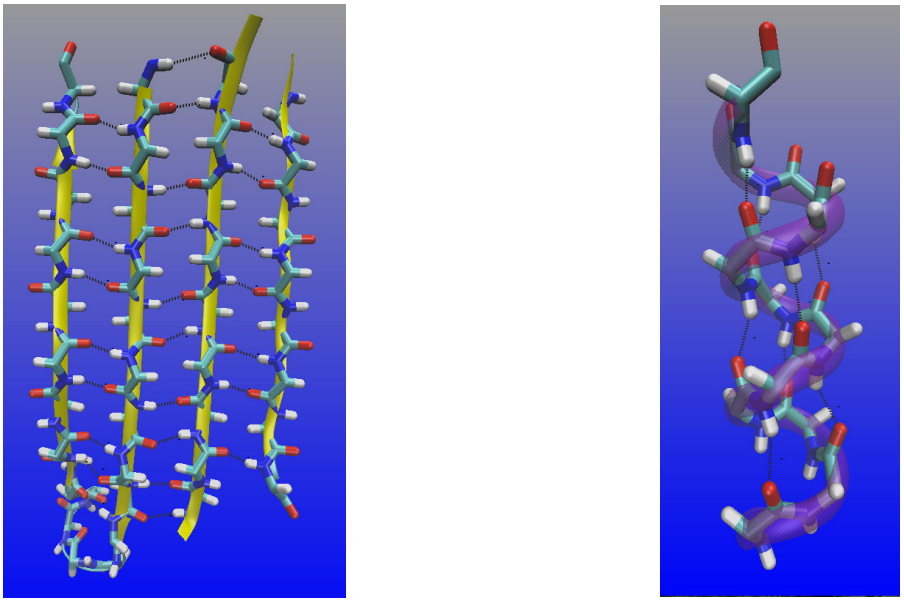


Fig. 4: *On the left the residues 48 to 55, 107 to 130 and 187 to 196 of concanavalin A (pdb code 1XQN) are shown as an example of an anti-parallel β-sheet [2] On the right residues 104 through 114 of sperm whale myoglobin (pdb code1L2K), are shown forming an α-Helix (Images made with VMD version 1.8.7).*

The side chains of the amino acids forming an α-Helix point to the outside perpendicular to the helix axis. Per winding 3.6 amino acids form one winding. The hydrogen bonding pattern can be seen in Fig. 4 or taken from Table 1. Since all hydrogen bonds have a dipole moment which is aligned in parallel in an α-helix a large dipole moment is formed by an α-helix which makes it energetically unfavourable when the number of residues involved exceeds 40 [4].

The β-sheet comes in two flavors a parallel one and an anti-parallel one (shown in Fig. 4). They differ by the hydrogen bonding pattern, but in both cases the side chains point roughly perpendicular to the plane defined by the backbone alternatingly upwards and downwards.

Table 1 Geometic properties of some secondary structure elements (values taken from [4]).

Secondary structure	Frequency	H-bonding	Handedness	Typical Φ	Typical Ψ
α-helix (3.6_{13})	abundant	i to i+4	right	-57°	-47°
3_{10} helix	infrequent	i to i+3	right	-20°	-54°
π-helix (4.4_{16})	rare	i to i+5	right	-57°	-80°
polyproline II	rare	-	left	-78°	+149°
polyclycine II	rare	i to i+3	left	-80°	+150°
parallel β-sheet	abundant	wide pair	-	-119°	+113°
antiparallel β-sheet	abundant	close pair	-	-139°	+135°

Apart from the structures mentioned in Table 1 certain turns form regular hydrogen bonding patterns such they can be considered secondary structural elements. On the bottom left of Fig.

4 with the backbone marked in blue a turn motif is visible. This is also an example of a secondary structure element of proteins. It allows for the amino acid chain to reverse its direction to form the anti-parallel β-sheet.

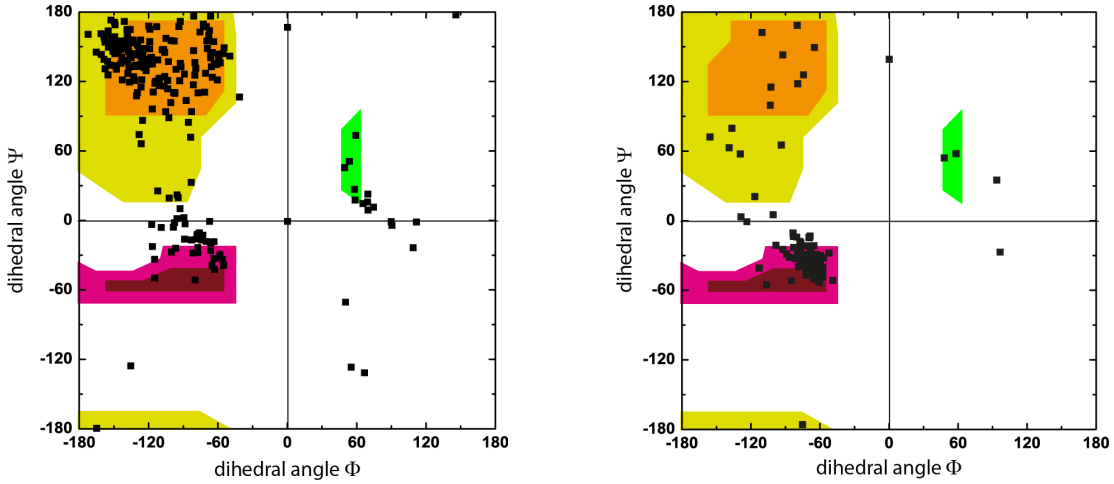


Fig. 5: *Ramachandran plots of the protein concanavalin A on the left and myoglobin on the right. The red area shows dihedral angles typical for α-helices. In the yellow and orange areas dihedral angles typical for β-sheets can be found. The green region corresponds to rare left handed helical arrangements of the protein backbone (Images made with VMD version 1.8.7 [5]).*

Ramachandran plots are especially suited to judge the secondary structure content of a protein. They consist of a scattered plot of all dihedral angles found per residue in the protein (see Fig. 5). On the x-axis all the dihedral angles of Φ and on the y-axis the dihedral angle Ψ is drawn for each amino acid residue resulting in one black filled symbol per residue. Obviously, myoglobin is mostly an α-helical and concanavalin A a β-sheet rich protein.Due to steric hindrances because of the presence of the side chain many combinations of Φ and Ψ are unfavourable (white areas in Fig. 5).

The **tertiary structure** of a protein denotes the three dimensional arrangement of all atoms of the protein in space, including the side chains. This information can be obtained by structural techniques like protein crystallography (will be discussed below), NMR. On that level all interactions mentioned above play a role. Supersecondary structural elements e. g. the α-helix bundle found in myoglobin and the βαβαβ-structure found in dehydrogenases are often seen in a tertiary structure. Another concept to divide the tertiary structure into sub-motifs is to define certain domains which are parts of the protein which can fold to this domain structure without the context of the complete protein. Often these domains provide functional sub-units and their structure is highly conserved throughout the protein family. Figure 6 shows three domains DI to DIII of human urokinase plasminogen activator receptor protein as an example.

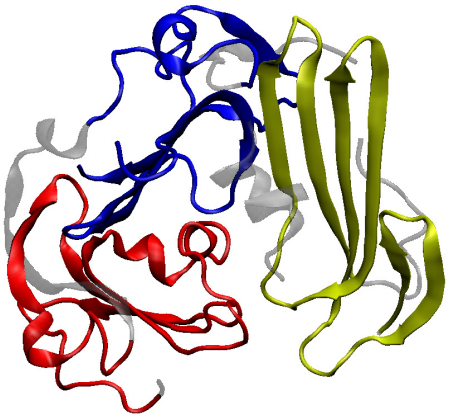


Fig. 6: *Secondary structure plot of human urokinase plasminogen activator receptor [6], a complex protein (pdb code 1YWH). The domains DI (yellow), DII (blue) and DIII (red) are shown. Amino acids not belonging to any domain are depicted in grey.*

Some proteins need more than one amino acid chain to be functional. The arrangement of the different amino acid chains is then referred to as the protein's **quarternary structure**.

2.3 The protein folding problem

The process with which the proteins reach this three dimensional structure is called protein folding and is under intense investigations after C. B. Anfinsen has performed pioneering experiments on denaturation and re-folding [7]. Considering the time a typical protein needs for folding which is of the order of seconds there must be some directive force leading to the correct fold. An exhaustive search of the overall parameter space of all possible dihedral angles Φ and Ψ would take too long for proteins of a typical size of 100 amino acids (Levinthal's paradox). A possible mechanism for such a directive force is the hydrophobic collapse where all hydrophobic side chains move together to from a hydrophobic core inside the protein. In a different hypothesis secondary structure elements of proteins form first and lead then to the final three dimensional fold.

3 Protein Crystallography

The previous chapter was intended to define some technical terms which describe protein structure in general. The following chapter will show how most of this structural information has been obtained.

The protein data bank (www.rcsb.org) is a well known source of structural information on proteins. Data from many different experimental techniques are entered in a standardized format, a .pdb file which essentially contains not only three dimensional coordinates of all observed atoms in a protein but also information on their mean square displacements. The number of stored entries exceeds 70 000. Among them x-ray crystallography has contributed more than 85 %. The next in line method with more than 8000 entries in the data bank is solution NMR spectroscopy. Electron microscopy as a method was used in more than 250 entries. The remaining methods count less than 100 entries each including neutron protein crystallography. Since the latter technique is represented by an instrument in the Jülich Centre

for Neutron Science (JCNS) and because of its similarity to x-ray crystallography it will be given some space in this lecture.

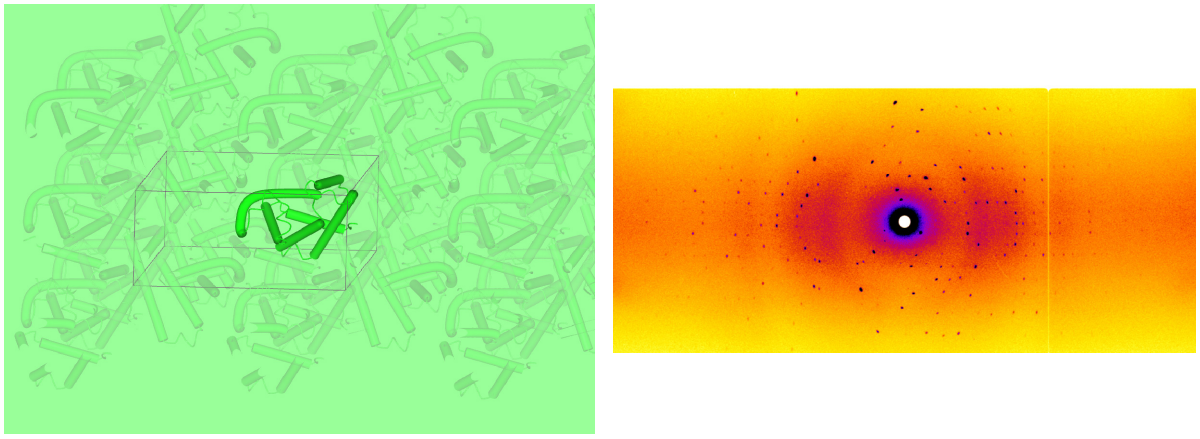


Fig. 7: *Real space arrangement of myoglobin molecules in a crystal of space group $P2_1$ (on the left) versus diffraction pattern (right) of a myoglobin crystal.*

For both techniques x-ray and neutron protein crystallography a single crystal of the protein of interest is required since the scattering of one protein molecule is very weak. Only using very bright x-ray sources in the future (e. g. XFEL) one might be able to gain enough information out of many single protein molecules in solution, averaging many exposures and orientations. But in general a crystal has to be grown, especially large ones in case of neutron crystallography since the neutron luminosity of modern sources is much smaller than the x-ray flux reached by synchrotron sources. To grow sufficiently large crystals is a big challenge in the case of many proteins, especially membrane proteins. Here, one has to adjust a large parameter set of protein concentration, pH condition, salt concentration, percipitant concentration and type just to name a few. The crystal then serves as a noiseless amplipier of the diffraction signal. But the arrangement of proteins in a crystals brings in another advantage, since the orientational averaging can be avoided, which is always present in the solution phase. Fig. 7 shows on the left the regular arrangement of myoglobin molecules in a crystal lattice. The unit cell of the monoclinic lattice (space group $P2_1$) is indicated by black lines. It bears two myoglobin molecules in one unit cell. The picture on the right shows a diffraction pattern recorded with the instrument BioDiff on a myoglobin crystal. The crystal is rotated by ca. 0.5° while recording one diffraction pattern. In order to map the reciprocal space completely one has to put the crystal in many different orientations into the beam and record diffraction patterns as mentioned above. Fortunately, crystal symmetry helps that some orientations are equivalent to each other and need not be recorded.

3.1 An X-ray protein crystallography beamline

Synchrotron beamlines provide extremely high photon flux for x-ray protein crystallography. Due to the high demand from the structural biology community, often more than one protein crystallography beamline is operated at a synchrotron. Those beamlines are optimized for special wavelengths and focal diameters. Here as an example the beamline BL14.2 is used to elucidate a typical x-ray protein crystallography experiment.

Figure 8 shows a schematic view of the beam path of beamline BL14.2. The beam paths to the other beamlines BL14.1 and BL14.3 have been omitted for clarity. A supraconducting magnet-structure interacts with the electron orbit of the storage ring and creates the so called synchrotron radiation. This radiation is used as a white light X-ray source for the three beamlines. A double crystal monochromator is used to select a very narrow energy band (2 eV at 9 keV) from the broad spectrum of the X-ray source. The mechanics of the double crystal monochromator keeps the out-going beam path constant when changing the wavelength. Focussing mirrors and collimators in the beam path ensure an efficient photon transport from the source to the sample and a small beam size at the sample position of 150 µm x 100 µm (hxv, FWHM). The Rayonix MX-225 detector has a pixel size of 37 µm. Without on chip binning one frame amounts to 6144 x 6144 pixels. The exposure time per frame is typically between 3 to 10 seconds.

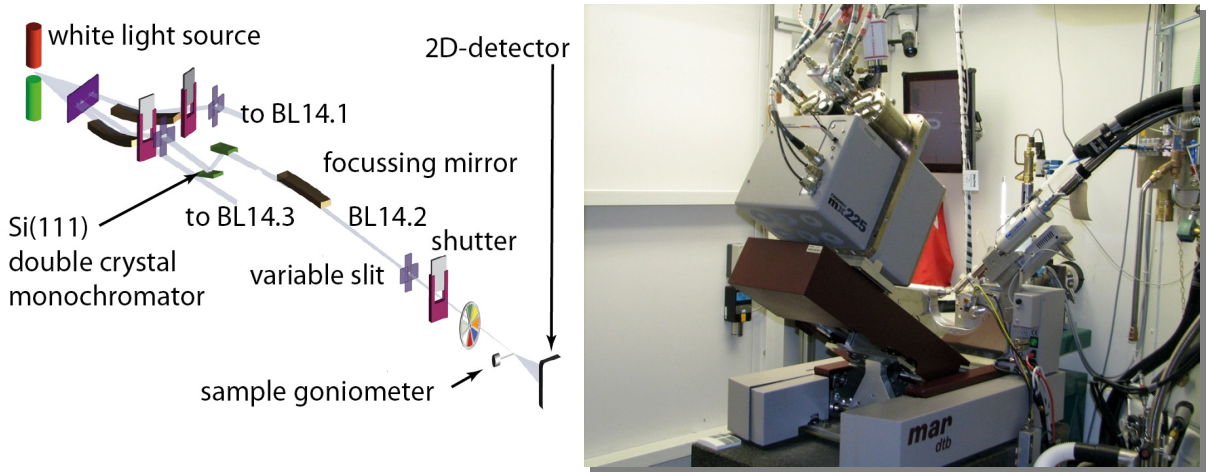


Fig. 8: *Schematic view of the beamline 14.2 of BESSY beginning with the magnet structure which is used as a white light x-ray photon source. On the right a picture taken from the experimental hutch of beamline 14.1 is shown. The beam enters from the right and the sample goniometer is mounted horizontally from the left. A cryostream sample environment to stabilize the sample temperature is discernible pointing towards the sample from the top right corner. Pictures kindly provided by Dr. Uwe Müller.*

Typically the sample crystals are kept at liquid nitrogen temperatures to avoid radiation damages. To record a full data set takes about 10-30 minutes. The largest diagonal of a typical protein crystal ranges between 10 to 500 µm

3.2 A neutron protein crystallography instrument

Since x-rays are scattered from the electrons in the crystal and neutrons from the nuclei, hydrogen atoms are hardly seen in x-ray crystallography experiments. Only at very high resolutions of 1 Å or less there is a chance to observe hydrogen atom positions. This resolution is often not within reach because of the crystal quality. Here neutron protein crystallography must be employed to retrieve the hydrogen atom positions. Moreover, neutron scattering can distinguish between different isotopes, especially between hydrogen and

deuterium. Whereas from x-ray crystallography the electron density in the unit cell of the crystal can be calculated, neutron protein crystallography yields the nuclear scattering length density, which is a signed quantity. In fact, the coherent scattering length of hydrogen is negative and the one from deuterium is positive (cf. Fig. 9).

Nucleus	atomic number	scattering length [10^{-12} cm]
^{1}H	1	-0.378
^{2}H	1	0.667
^{12}C	6	0.665
^{15}N	7	0.921
^{16}O	8	0.581

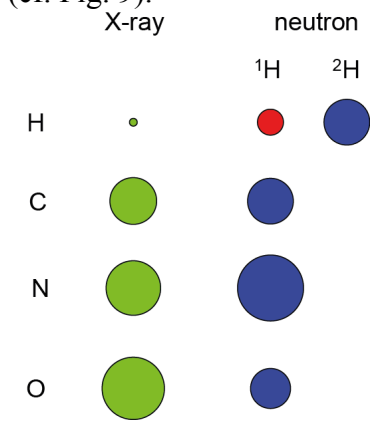


Fig. 9: *The table on the left lists scattering lengths of selected atoms of biological relevance. On the right there is a comparison of x-ray scattering cross section with scattering lengths from neutron scattering. The circles are scaled to match at the carbon atom.*

A major drawback of the method neutron protein crystallography is the required crystal size. Due to the much smaller neutron flux as compared to x-ray flux the crystals required for a neutron crystallography study must be much larger as compared to x-ray crystallography. Here, often crystal diagonals of 1 mm and more have to be reached.
As an example of a neutron diffraction instrument optimised for protein crystallography the instrument BioDiff at the FRM II shall be introduced. It is a collaboration between the Forschungszentrum Jülich (FZJ) and the Forschungs-Neutronenquelle Heinz Maier-Leibnitz (FRM II).

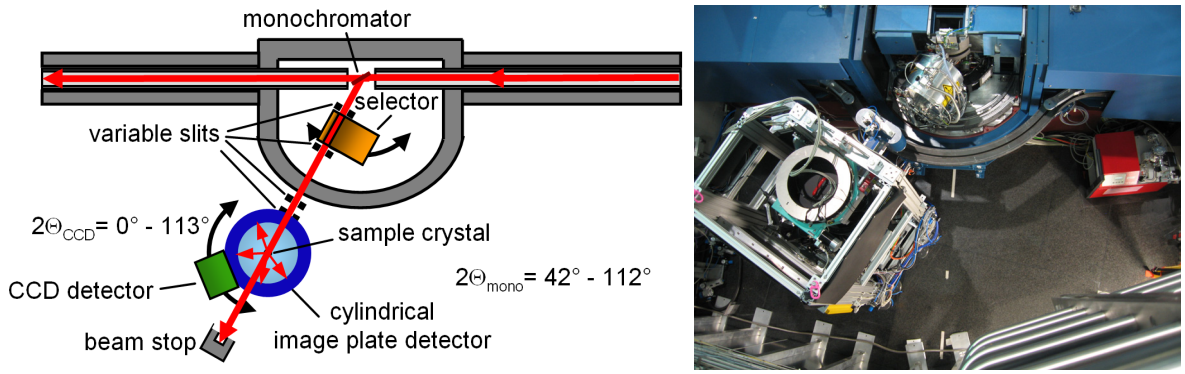


Fig. 10: *Schematic view of the BioDiff instrument (left) and a picture taken from a similar view point with the biological shielding removed (right).*

Figure 10 shows a schematic view of the instrument from the top and a corresponding picture when the biological shielding has been removed. The neutron beam from the cold source of

the FRM II reactor enters from the right. By Bragg reflection from a pyrolytic graphite crystal (002-reflex) neutrons are taken out from the white neutron spectrum of the neutron guide NL1 and pass a first boron carbide adjustable slit and a velocity selector. The velocity selector acts as a $\lambda/2$ filter. Together with the pyrolytic graphite crystal it forms a monochromator with a $\Delta\lambda/\lambda$ of ca. 2.5 %. Behind the velocity selector the beam passes a second variable slit and the main instrument shutter, named γ-shutter. Additionally, a boron carbide neutron shutter is placed directly after the monochromator crystal for a faster shutter operation. Before entering the detector drum of the image plate detector through a Zirconium window a collimator made out of two manually exchangeable boron carbide apertures with fixed diameters between 3 mm and 5 mm shape the beam to fit to the sample size. At present the sample is usually contained in a glass tube (either a thin walled capillary or a NMR-tube for larger crystals). It is fixed to a standard goniometer which is mounted upside-down from the sample stage on top of the instrument. After passing the sample the main neutron beam exits the detector drum through a second Zirconium window and hits finally the beam stop which consists of a cavity of 4 mm thick boron carbide plates surrounded by a 13 cm thick wall of lead shielding bricks. The cylindrical image plate detector is covering roughly half of the total 4π solid angle. It is 200 mm in diameter and 450 mm in height. It can be read out with three different resolutions of 125 μm, 250 μm and 500 μm. As an alternative, one can lower the image plate detector and swing in a neutron sensitive scintillator which is imaged onto a CCD-chip. This CCD-camera set up serves as a second detector. In particular it is used for a fast alignment of the sample crystal with respect to the neutron beam.

With the image plate detector the diffraction pattern shown in Fig. 7 on the right has been recorded. In fact, a complete crystallographic data set on a myoglobin crystal has been recorded allowing for the calculation of a nuclear scattering length density map. The exposure time was 17 minutes per frame and the crystal was rotated by 0.5° during exposure. 331 frames were recorded before the crystal was manually rotated by ca. 90° in the capillary and another set of 243 frames were recorded. Altogether ca. 9 days of beam time were necessary to record the complete data set. The achieved resolution with sufficient completeness was 1.7 Å. The required time to record this data set was much longer as the 30 minutes from x-ray diffraction.

3.3 Some general aspects of diffraction by a protein crystal

Having recorded a complete data set on a crystal some data treatment is necessary in order to calculate meaningful atom positions. Here only a brief introduction can be given more details can be found in text books [8-9].
Assuming a number of n atoms per unit cell the structure factor of a single unit cell can be written as (see previous lectures):

$$F(\mathbf{S}) = \sum_{j=1}^{n} f_j \exp(2\pi i \mathbf{r}_j \mathbf{S}) \qquad (1)$$

Here $\mathbf{r}_j$ denote the atom position of atom j and $\mathbf{S}$ is the scattering vector perpendicular to the plane which reflects the incident beam. f_j can be seen here either as the scattering length of atom j in the neutron scattering case or the atomic scattering factor in case of x-ray diffraction. One can generalize this approach by switching form the summation to an integration to yield:

$$F(\mathbf{S}) = \int_{unitcell} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r}\mathbf{S}) d^3\mathbf{r} \tag{2}$$

where $\rho(\mathbf{r})$ is the electron density distribution or the scattering length density respectively. Since a crystal consists of AxBxC unit cells, the structure factor of the crystal can be composed as

$$F_{cryst.}(\mathbf{S}) = F(\mathbf{S}) \sum_{u=0}^{A} \exp(2\pi i u \mathbf{a}\mathbf{S}) \sum_{v=0}^{B} \exp(2\pi i v \mathbf{b}\mathbf{S}) \sum_{w=0}^{C} \exp(2\pi i w \mathbf{c}\mathbf{S}) \tag{3}$$

The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ denote basis vectors of the unit cell. For an increasing number of unit cells the sums can be represented by delta functions leading to the Laue conditions for the structure factor being non-zero:

$$\mathbf{a}\mathbf{S} = h, \mathbf{b}\mathbf{S} = k, \mathbf{c}\mathbf{S} = l \tag{4}$$

This means that one only gets constructive interference, when the scattering vector is perpendicular to planes in the crystal which can be denoted by the Miller index vector $\mathbf{h} = hkl$. For this reason the diffraction pattern of a single crystals shows distinct peaks, the so called Bragg peaks. The Bragg law can be easily derived from equation 4. Figure 11 shows the Ewald sphere construction. It is a tool to construct the direction of the diffracted beam. The Ewald sphere has its origin at the position of the crystal. Its radius is the reciprocal wavelength used in the scattering experiment. The origin of the reciprocal space lattice is placed at the intersection of the sphere with the incident beam direction. Whenever the orientation of the reciprocal space is such that another point of the reciprocal space lies on the Ewald sphere a diffracted beam results in the direction of line running from the centre of the Ewald sphere through that point.

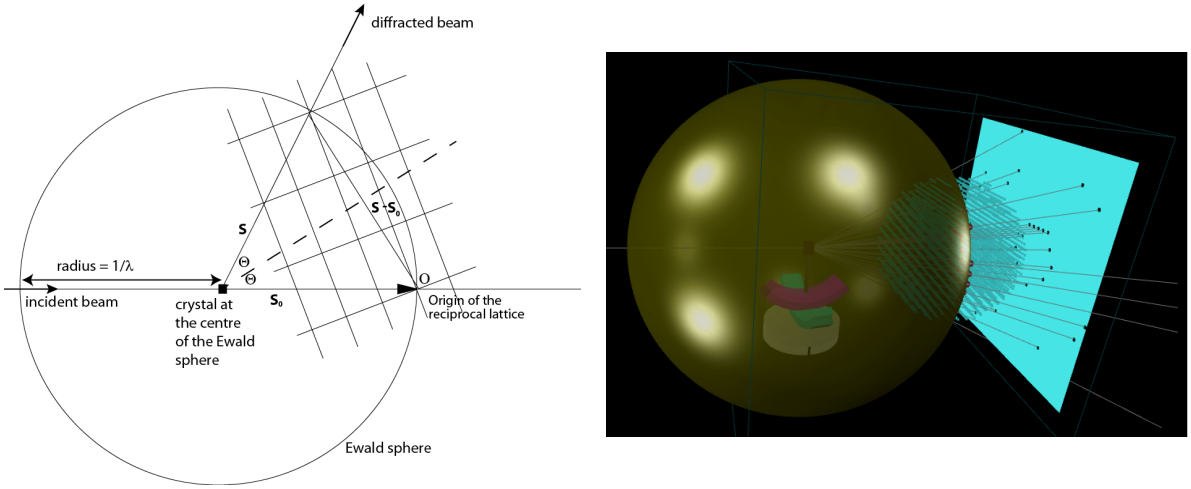


Fig. 11: *Ewald sphere: On the left the schematic shows how to construct the Ewald sphere. On the right an Ewald sphere (golden colour) construction is shown in three dimensions. The blue square represents a flat two dimensional detector.*

When the crystal is rotated the reciprocal space rotates with it resulting in other lattice points to cause diffracted beams. In practice the incident beam is not strictly monochromatic but has a wavelength distribution which causes the Ewald sphere to be elongated to form a spherical shell of a certain thickness. This increases the number of diffracted beams observed. The beam divergence adds also to its thickness.
So, the positions of the diffracted beams on the detector only depend on the reciprocal lattice. The structure of the protein inside the unit cell is encoded in the amplitude and phase of the structure factor.
To obtain the electron density or the nuclear scattering length density one has to perform the inverse Fourier transformation:

$$\rho(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{h}} F(\mathbf{h}) \exp(-2\pi i \mathbf{rh}) \qquad (5)$$

Here V is the volume of the unit cell. Unfortunately only the modulus squared of the structure factor is measured as intensity on the detector. The phase information is lost which is known as the phase problem of crystallography.

There are several solutions to the phase problem which are only applicable for the x-ray diffraction case:

- isomorphous replacement: Several crystals of the same crystal structure have to be available for this method. First a crystallographic data set is recorded on an untreated crystal. Then crystals are soaked in at least two different heavy atom salt solutions. In the best case, the different heavy atom ions occupy different regular positions in the unit cell. From these (at least) two crystallographic data sets recorded on the heavy atom treated crystals phase information can be retrieved which is then used to determine the phases of the data set of the untreated crystal.

- anomalous dispersion: Often it is possible to replace one distinct methionine amino acid with an artificial selenomethionine one. The selenium atom has a suitable absorption edge on which anomalous scattering can be performed by tuning the wavelength of the beamline to the anomalous regime. Crystallographic data sets are then recorded at different wavelengths from with the phase information can be calculated. In some cases this approach can also be adopted for sulfur atoms present in naturally occuring cysteine residues.
- molecular replacement: From the primary structure one can search the protein data base (pdb) for proteins with a similar amino acid sequence. If one finds enough fragments which seem to be sufficiently homologous to the unknown structure one can use those fragments for the calculation of initial phases. In further refinement steps these phases can be improved further. Since the number of unique structures entered in the protein data base is growing this method is increasingly favoured over other methods.

The phase problem of the neutron data sets is solved by using the x-ray structure and the molecular replacement technique.

3.4 Model building and refinement

With the data treatment one has now arrived at a contour map $\rho(\mathbf{r})$ be it either a nuclear scattering length density or an electron density. Now the information on the primary structure of the protein is used and either manually or employing software first the backbone is coarsely fitted into the contour map. Then from this model new amplitudes and phases of the structure factor are calculated using eq. 1. The modulus squared of the structure factor is again compared with the data. One could now think of a least square based fitting procedure to find the optimum arrangement of the protein atoms in the unit cell. In practice however maximum likelihood and simulated annealing molecular dynamics simulations are used because those are superior to the least square approach in terms of overcoming local minima. In these molecular dynamics simulations a lot of stereochemical information is used as restraints for example known bond lengths of CC single bonds or bond angles. The agreement between model and observed contour map is often measured by calculating a so called R-factor:

$$R = \frac{\sum_{\mathbf{h}} \left| \left| F_{obs}(\mathbf{h}) \right| - \left| F_{calc}(\mathbf{h}) \right| \right|}{\sum_{\mathbf{h}} \left| F_{obs}(\mathbf{h}) \right|} \tag{6}$$

The index "obs" denotes the observed structure factors and the index "calc" the calculated structure factors from the model. The value of the R-factor lies in the limits between 0 and 1. A good agreement between model and measured data leads to an R-factor of about 0.2. R-factors of 0.5 and above are indicative for a random agreement between model and data.

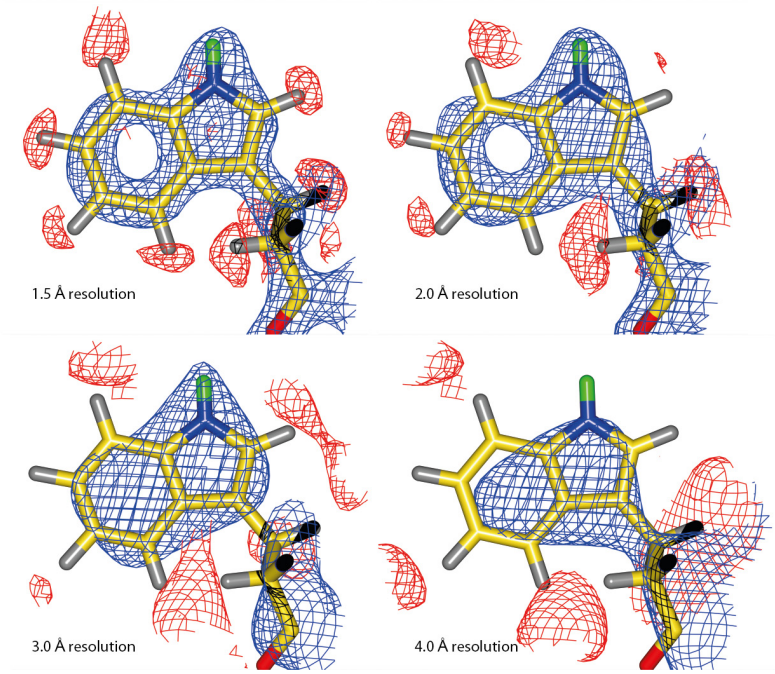


Fig. 12: *The side chain of the amino acid tryptophan no. 7 of myoglobin measured with neutron diffraction at different resolutions. The contour level of the shown nuclear map is +1.5σ (blue) and -1.75σ (red). Exchanged (liable) hydrogen atoms (green) and C- (yellow), N- (blue) and O-atoms (red) appear on a positive contour level. Only not liable hydrogen atoms are seen on the negative contour level.*

But even a good R-factor does not guarantee that the model fits to the data. In fact, it is possible in special cases to obtain a reasonably low R-factor when using the amino acid chain in the wrong direction as a model [10]. Here, Brünger et al. [11] have suggested to divide the measured Bragg reflections into two subset one working set denoted by "A" and one test set denoted by "T". With the working set the fitting procedure is performed, whereas the test set only serves to control the model quality by calculating the R_{free} factor.

$$R_{free} = \frac{\sum_{\mathbf{h}\in T} \left| |F_{obs}(\mathbf{h})| - |F_{calc}(\mathbf{h})| \right|}{\sum_{\mathbf{h}\in T} |F_{obs}(\mathbf{h})|} \qquad (7)$$

The test set usually consists of 5 to 10 % of all structure factors, uniformly distributed over the resolution range.

The R-factor and R_{free} factor should not differ too much from each other. In general, it is good practice to always look at the resulting model and its fit to the calculated map after each refinement step. Ramachandran plots can also be used to judge whether the amino acid backbone adopts a reasonable fold. With decreasing resolution (cf. Fig. 12) of the data it becomes more and more difficult to find the right orientations of side chains or even errors in

the registry of the protein backbone can occur, whereby for example one amino acid is left out.

4 A case study: Water network around myoglobin

This example is chosen since it nicely shows the interplay between x-ray and neutron crystallography. Myoglobin has been used quite frequently as an example throughout this lecture. Its function is to take over the oxygen molecules from the blood heme proteins in the red blood cells and to transport and store it within the muscle cells. Therefore its binding affinity to oxygene must be stronger than that of the hemoglobin. In order to perform the transport tasks myoglobin has to be highly soluble and movable within the context of a muscle cell. Let alone therefore it is interesting to study the water network surrounding of myoglobin. Since it is a joint neutron and x-ray diffraction study the crystal under investigation has been soaked in D_2O in order to exchange all liable hydrogen atoms with deuterium atoms.

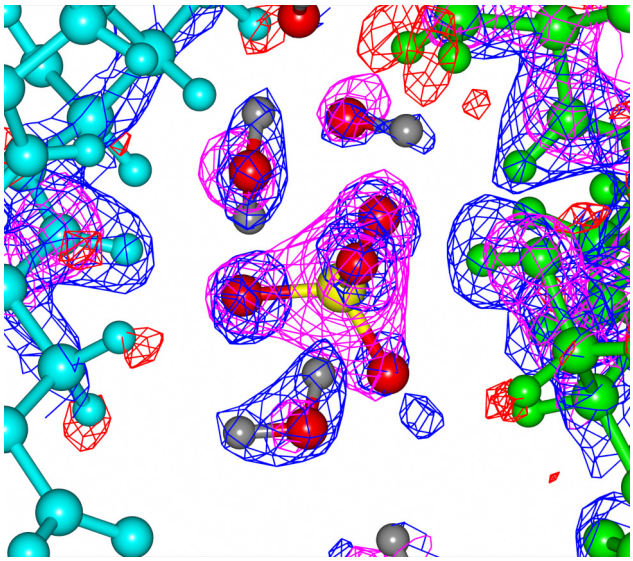


Fig. 13: *Water network in the contact region between two myoglobin molecules in the crystal. In grey colour on the left amino acids 51 to 52 of one myoglobin molecule are shown. On the right amino acids 58 to 60 (from top to bottom) are depicted in green. In the centre of picture a sulfate ion SO_4^- is seen with the sulfur atom shown in yellow, the oxygen atoms shown in red. The deuterium atoms of the water molecules are coloured grey. The x-ray map is shown in magenta at a contour level of +2.7s. The nuclear map is shown at a contour level of -1.75s in red and at +2.3s in blue. Data taken from ref. [3]. (The picture is similar to the one shown in [12] or [13] since it is based on the same data.)*

Figure 13 shows a contact region between two myoglobin molecules. In the centre a sulfate ion is observed. Here, in the nuclear map the central sulfur atom is hardly seen because of its small scattering length. The oxygen atoms of the sulfate ion and of the water molecules are

readily observed by both techniques. The deuterium atoms of the water molecules are discernible only in the nuclear map. When the water molecule is fixed by hydrogen bonds, all three atoms can be observed. These water molecules exhibit a triangular shape in the nuclear map. In case it is free to rotate around the OH-axis the nuclear scattering length of the rotating deuterium atom is distributed over a large volume and is therefore averaged out. Those water molecules are denoted as "short ellipsoidal" by Chatake et al. [12]. The long ellipsoidal water molecules are fixed at both deuterium atoms but only the oxygen can rotate freely around the DD-axis. In the fourth case only the oxygen atom is observed. In this case the orientation of the water molecule is not fixed, only the oxygen atom is held in place.

This classification of water molecules helps to judge the flexibility of the water network around proteins. It can also be used to classify observed water molecules and their hydrogen bonding pattern in molecular crystals in general.

Acknowledgement

The author would like to thank Christian Felder for software support, Dr. Weiergräber for providing his lecture notes and transparencies of the 2011 spring school and Prof. Georg Büldt for valuable comments. The author is indebted to Dr. Uwe Müller from BESSY for his timely help on the X-ray beamline chapter. The author is especially grateful to Dr. Andreas Ostermann for supplying some of the Figures and for carefully reading the manuscript.

References

[1] L. Stryer, *Biochemie* (Spektrum Akad. Verlag, Heidelberg Berlin New York, 1991).
[2] M. P. Blakeley *et al.*, Proc. Natl. Acad. Sci. U. S. A. **101**, 16405 (2004).
[3] A. Ostermann *et al.*, Biophys. Chem. **95**, 183 (2002).
[4] O. H. Weiergräber, in *Macromolecular Systems in Soft and Living Matter*, edited by J. K. G. Dohnt *et al.* (Forschungszentrum Jülich GmbH, Institute of Complex Systems, Jülich, 2011).
[5] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).
[6] P. Llinas *et al.*, EMBO J. **24**, 1655 (2005).
[7] C. B. Anfinsen, Comparative Biochemistry and Physiology **4**, 229 (1962).
[8] J. Drenth, *Principles of Protein X-Ray Crystallography* (Springer Science+Business Media, LLC, New York, 2007), p. 332.
[9] N. Niimura, and A. Podjarny, *Neutron Protein Crystallography - Hydrogen, Protons, and Hydration in Bio-macromolecules* (Oxford University Press, Oxford, New York, 2011).
[10] G. J. Kleywegt, and T. A. Jones, Structure **3**, 535 (1995).
[11] A. T. Brunger, Nature **355**, 472 (1992).
[12] T. Chatake *et al.*, Proteins-Structure Function and Genetics **50**, 516 (2003).
[13] T. Chatake *et al.*, Journal of Synchrotron Radiation **11**, 72 (2004).

Recommended Textbooks

on proteins in gerneral:

L. Stryer, *Biochemie* (Spektrum Akad. Verlag, Heidelberg Berlin New York, 1991)

on X-ray crystallography:

J. Drenth, *Principles of Protein X-Ray Crystallography* (Springer Science+Business Media, LLC, New York, 2007)

on neutron protein crystallography:

N. Niimura, and A. Podjarny, *Neutron Protein Crystallography - Hydrogen, Protons, and Hydration in Bio-macromolecules* (Oxford University Press, Oxford, New York, 2011)

F 1 Advanced Transmission Electron Microscopy Techniques and Applications

R.E. Dunin-Borkowski, M. Feuerbacher, M. Heggen, L. Houben,
A. Kovács, M. Luysberg, A. Thust and K. Tillmann
Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons
Forschungszentrum Jülich GmbH

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research:
Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

Modern transmission electron microscopes (TEMs) can be used to obtain quantitative measurements of the structural, electronic and chemical properties of materials on length scales down to the sub-Å level [1-3]. In the last decade, electron microscopy has been revolutionized by the introduction of aberration correctors, monochromators and imaging filters, as well as by improvements in computing power for microscope control, image analysis and image simulation. Specialized techniques have also been developed, including the use of *in situ* gas reaction electron microscopy to study growth processes and chemical reactions in materials at elevated temperature and pressure and the application of bright and dark field electron holography to measure variations in electrostatic potential, magnetic induction and crystallographic strain in materials with sub-5-nm spatial resolution. A photograph of a modern aberration corrected transmission electron microscope is shown in Fig. 1.

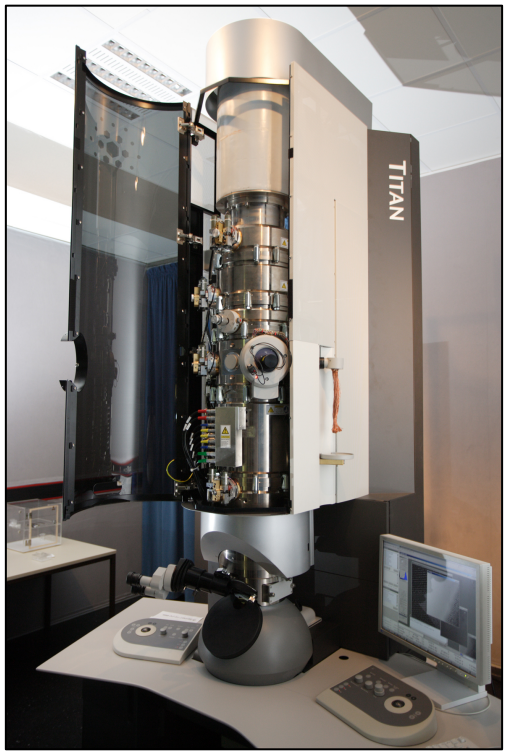


Fig. 1: *FEI Titan 80-300 field emission transmission electron microscope at Forschungszentrum Jülich. The instrument is equipped with a spherical aberration corrector on the objective lens and has an information limit of 0.08 nm*

The basic operation of a TEM is in many respects analogous to that of a light optical microscope. The electron source at the top of the column is either a thermionic emitter such as a LaB_6 single crystal or a field emitter such as a zirconia-coated tungsten tip. The primary advantage of a field emission source is that it is brighter, as a result of the smaller extraction area for electrons, which can be only a few nm in size. After emission, the electrons are accelerated by a voltage of typically between 80 and 300 kV and focused onto the specimen

by 2-3 condenser lenses. The specimen is most commonly a 3 mm disk that is prepared to be extremely thin at its centre, as the typical specimen thickness for high-resolution imaging is below 10 nm. The specimen stage allows movement of the sample in three spatial directions and tilting about two axes. The objective lens is located directly below the specimen and has a short focal length of 1-2 mm. Its design and the stability of its power supply are crucial for the optical performance of the microscope. The intermediate and projector lenses in the lower part of the column are used for magnification of the image and have a relatively small influence on image quality. After passing the projector lens, the electrons can be observed on a fluorescent screen or recorded on a CCD (charge coupled device) camera. Active vibration damping systems and electromagnetic field compensation systems are normally required to create a sufficiently stable environment for a state of the art electron microscope.

2 Aberration correction in high-resolution TEM

High-resolution transmission electron microscopy (HRTEM) involves the acquisition of images in a TEM with a spatial resolution that is sufficient to separate single atomic columns. The interpretation of such images is, however, not straightforward, as the recorded intensity is not a direct representation of the specimen but an interference pattern that is affected both by the strength of the interaction of the incident electrons with the specimen and by the contrast transfer of the microscope [4].

The interaction of an incoming electron wave with a TEM specimen can be calculated by solving the relativistically corrected Schrödinger equation for the electron wavefunction $\Psi(\mathbf{r})$ in a periodic crystal potential $V(\mathbf{r})$ according to the Bethe-Bloch formalism [5]. However, in practice, several approximations are often used to understand the image formation process. In the phase object approximation (POA) for a thin specimen, atoms in the sample are simulated by a projected potential that is continuous and constant in the direction of the incident beam. The electron wavefunction at specimen thickness t is then written in the form

$$\Psi(\mathbf{r},t) \approx \exp[i\Phi(\mathbf{r},t)] = \exp[i\sigma V_P(\mathbf{r},t)] \ , \tag{1}$$

where Φ is the phase of the electron wave, σ is an interaction constant and V_P is the projected crystal potential. The weak phase object approximation (WPOA) assumes that the phase modulation of the electron wave is small in a very thin specimen, resulting in the expression

$$\Psi(\mathbf{r},t) \approx 1 + \mathrm{i}\Phi(\mathbf{r},t) = 1 + \mathrm{i}\sigma V_p(\mathbf{r})t \tag{2}$$

The effect of the microscope lenses is described by modifying the exit plane wavefunction using a phase factor $\exp[-\mathrm{i}\chi]$ in the form

$$\Psi_{\mathrm{i}}(\mathbf{g}) = \Psi(\mathbf{g})\exp[-\mathrm{i}\,\chi(\mathbf{g})] \tag{3}$$

where, for the rotationally symmetric aberrations of a round lens, the aberration function

$$\chi(g) = 2\pi\left(\frac{1}{4}C_S\lambda^3 g^4 + \frac{1}{2}Z\lambda g^2\right) \quad (4)$$

In Eq. 4, λ is the wavelength of the electron beam, Z is the defocus and C_S is the coefficient of spherical aberration of the objective lens, which cannot be avoided for a round electromagnetic lens [6] and describes the deviation of rays passing the outer part of a lens compared to near-axis rays and the resultant blurring of the object, as shown in Fig. 2.

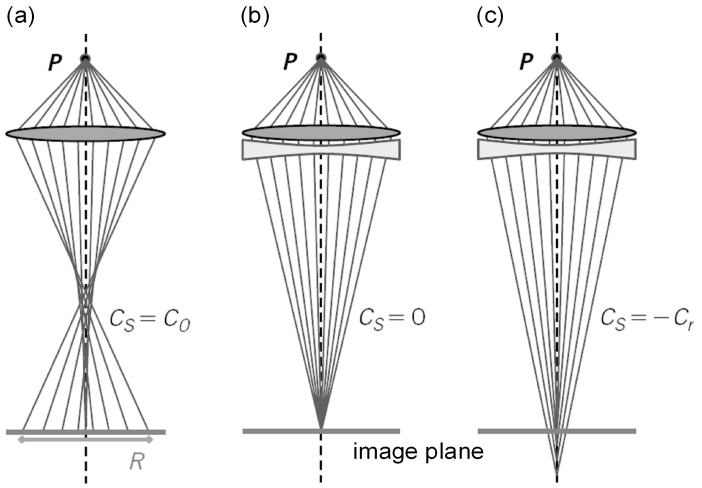


Fig. 2: *Schematic illustration of (a) positive spherical aberration, (b) no spherical aberration and (c) negative spherical aberration.*

For a thin specimen, the WPOA results in an expression for the linear image intensity of the form

$$I_L(\mathbf{g} \neq 0) \approx 2\sigma V(\mathbf{g}) t \sin\chi(\mathbf{g}) \quad , \quad (5)$$

from which it is apparent that the optimum image contrast is obtained when the coherent contrast transfer function (CTF) $\sin\chi = \pm 1$, i.e., when $\chi(\mathbf{g})$ is an odd multiple of $\pi/2$ for all values of g. However, $\sin\chi$ is in general a function that oscillates strongly with g. As a result, atomic columns that are arranged periodically with a spacing of $d = 1/g$ are imaged as white dots only at selected spatial frequencies for which the CTF is close to 1. At spatial frequencies for which the CTF is close to –1, the atomic columns are imaged as black dots, while they may be invisible if the spatial frequency corresponding to their interatomic spacing coincides with a zero of the CTF. In order to approach ideal phase contrast transfer behaviour, a defocus setting can be chosen that balances the g^2 and g^4 terms in the aberration function,

allowing for a relatively broad band of frequencies to be transferred with a CTF close to −1. This defocus setting

$$Z_S \approx -\sqrt{\frac{4}{3} C_S \lambda} \ . \tag{6}$$

is known as Scherzer defocus [7]. The point resolution $d_S = 1/g_S$ is defined by the first zero crossing of the CTF at the Scherzer defocus and is given by the expression

$$g_S \approx \left(\frac{3}{16} C_S \lambda^3 \right)^{-\frac{1}{4}} \ . \tag{7}$$

Typical values of the point resolution of commercially available medium voltage microscopes with accelerating voltages of 200-400 kV are in the range 0.24 to 0.17 nm.

The limited coherence of the electron source and electronic instabilities have the additional effect of multiplying the coherent CTF by envelope functions, which result in a cut-off of the contrast transfer at high spatial frequencies. The spatial frequency at which the partially coherent CTF falls below a threshold value defines the information limit of the microscope. For an uncorrected field emission TEM, the information limit can be higher than the point resolution, leading to strong oscillations in the partially coherent CTF and blurring effects in images. Figure 3 shows $\sin\chi$ for a 200 kV instrument for two different values of C_S. At higher spatial frequencies the CTF oscillates rapidly up to the information limit. If C_S is reduced then a broad transfer band extends up to the information limit of 0.125 nm, improving the point resolution significantly.

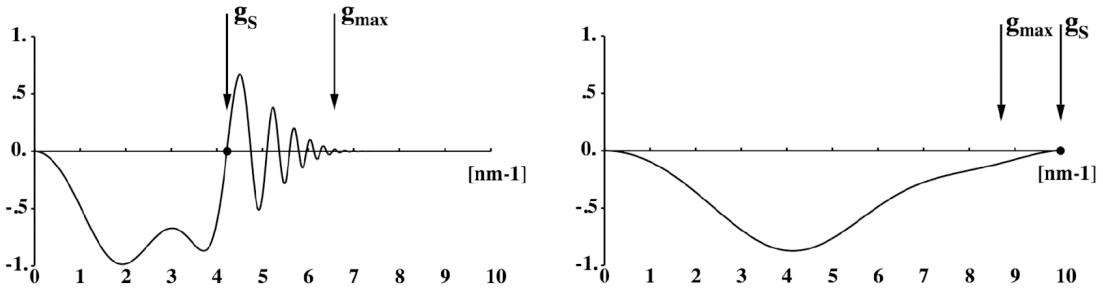


Fig. 3: *Partially coherent contrast transfer functions at Scherzer defocus conditions calculated for a CM20 FEG microscope operated at 200 kV for (left) C_S = 1.2 mm, and (right) C_S = 0.04 mm.*

An important approach that can be used to overcome the effect of the strongly oscillating part of the CTF between the point resolution and the information limit for an uncorrected microscope is to use numerical reconstruction of the exit plane wavefunction from a defocus

series of 10–20 images acquired from the same object area [8-11], as shown schematically in Fig. 5. According to the phase object approximation, the heights of the phase maxima in the wavefunction are approximately proportional to the projected crystal potential for a thin specimen, permitting the chemical distinction of atomic species. The availability of the complex-valued wavefunction also enables complete correction of all aberrations in software. The resulting wavefunction is then free from artefacts introduced by the microscope, making it possible to resolve light elements such as C, N and O close to or even below 0.1 nm spatial resolution.

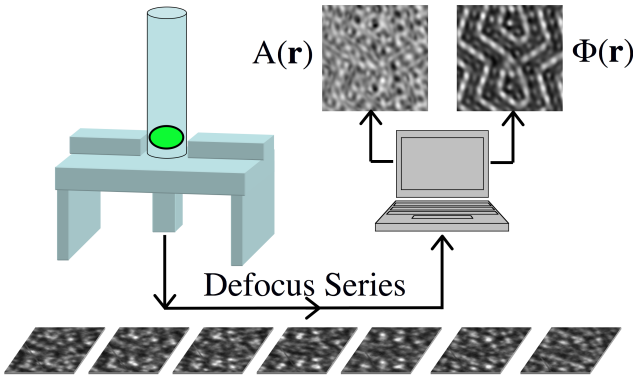


Fig. 4: *Principle of focal series restoration. A series of images is recorded from the same object area using different settings of the objective lens defocus. The quantum-mechanical exit plane wavefunction, which consists of amplitude and phase, can be retrieved from the series by means of numerical procedures.*

The elimination of delocalisation by focal series restoration using non-linear reconstruction from 20 images is shown in Fig. 5 for a twin boundary in $BaTiO_3$. The left side of Fig. 5 shows one image from the focal series, which was taken with an uncorrected Philips CM20ST FEG microscope. Blurring of the image due to objective lens aberrations is visible. The right side of Fig. 5 shows the phase of the restored wavefunction, in which blurring is no longer present and all atomic positions are resolved up to the information limit of the microscope.

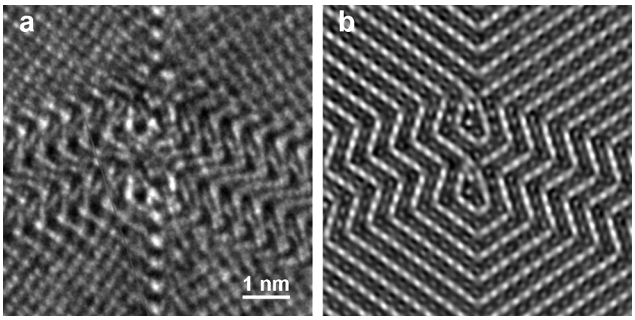


Fig. 5: *Example of focal series restoration. (a) Representative image from a series of 20 images of twinned [110] oriented $BaTiO_3$. (b) Phase of restored exit plane wavefunction. All aberrations have been removed numerically. Positions of atomic columns are revealed by phase maxima.*

The elimination of spherical aberration using hardware is now also possible using multipole lenses. The first successful demonstration of spherical aberration correction was achieved in the late 1990s in a project funded by the Volkswagenstiftung involving Forschungszentrum Jülich [12, 13]. By tuning the spherical aberration coefficient and other higher order aberrations, optimum contrast transfer and a dramatic improvement in resolution can be achieved. Figure 6 shows a comparison between an experimental defocus series of [110] $SrTiO_3$ recorded using a C_S corrected Philips CM200ST FEG microscope and simulated images. The structure image predicted by simulations for a defocus of +10 nm appears in the experiment.

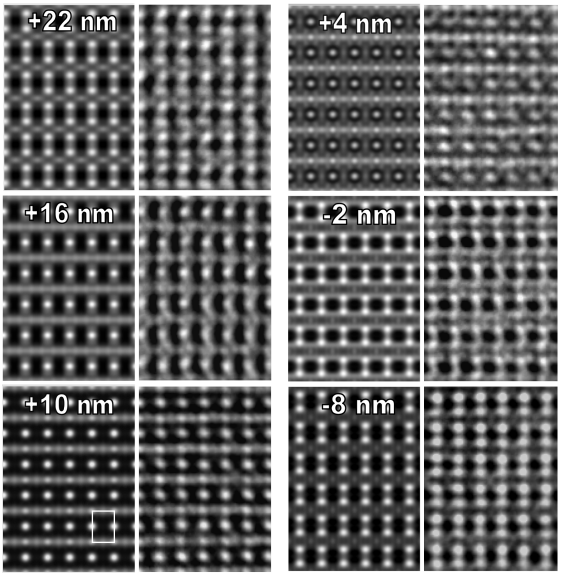


Fig. 6: *Comparison of an experimental defocus series of $SrTiO_3$ oriented along [110] with simulated images. The experimental images were acquired with a C_S corrected CM200ST FEG microscope using a value for C_S of –0.04 mm. Simulated images are shown on the left, with experimental images on the right. The defocus values of the simulated images are indicated. The specimen thickness used in the simulations was 3.5 nm. Half of a $SrTiO_3$ unit cell projected is indicated by the white frame.*

The ability to tune the spherical aberration coefficient opens up new ways to record directly interpretable images. In particular, the choice of a small negative value for C_S results in strong contrast from light atomic columns, which appear bright [14]. In this negative C_S imaging (NCSI) technique, paraxial rays travelling through the objective lens come to focus ahead of the outer rays and light elements such as oxygen in $Pb(Zr_{0.2}Ti_{0.8})O_3$ adjacent to columns of heavy atoms can be resolved [15].

It should be noted that the complete aberration function is made up of many individual aberrations. Sub-Å imaging requires both a knowledge of all significant wave aberrations up to sixth order in *g* and the stability of these aberrations over the length of an experiment [16]. The decomposition of a typical aberration function into its components is illustrated in Fig. 7, while the time stability of the defocus measured for two microscopes is show in Fig. 8.

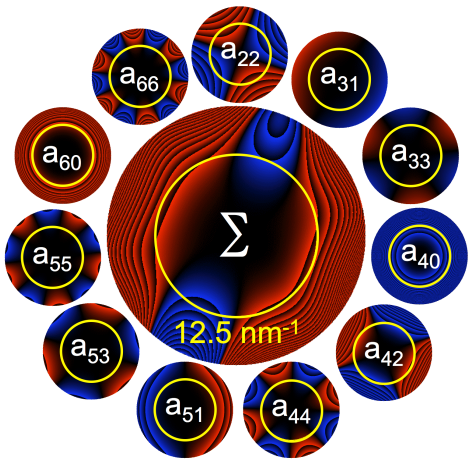


Fig. 7: *Aberration function (centre) and its decomposition into basic aberrations. Positive values of the aberration function are depicted in red and negative values in blue. Sawtooth jumps occur at intervals of π/4. The yellow circle marks a value of g = 12.5 nm^{-1}, corresponding to a resolution of 0.08 nm. Aberration-free imaging is achieved when the modulus of the total aberration function does not exceed π/4 within the circle defining the resolution limit.*

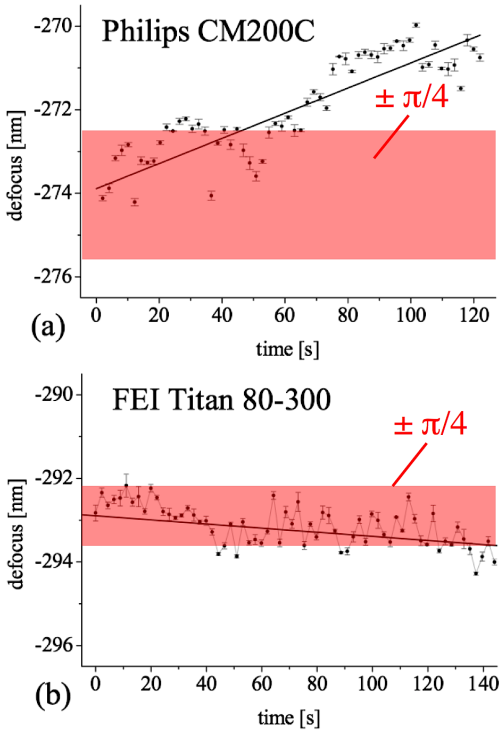
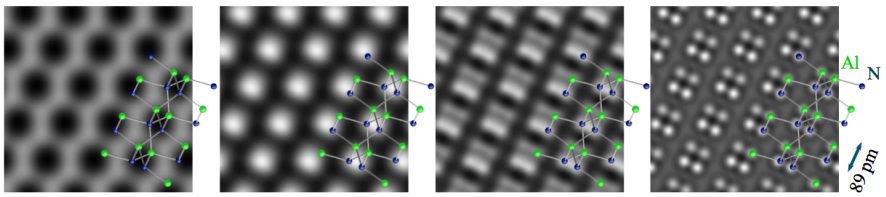


Fig. 8: *Time variation of the objective lens defocus during a period of approximately 2 minutes for (a) a 200 kV Philips CM200 microscope with a resolution of 0.12 nm and (b) a 300 kV FEI Titan microscope with a resolution of 0.08 nm. Both a thermally induced long-term drift and short-term fluctuations caused by instabilities of the accelerating voltage and the objective lens current are observed.*

A graphical illustration of the dramatic improvement in image resolution and interpretability with microscope generation and the use of spherical aberration correction is shown in Fig. 9 for simulated images of AlN.



CM30-UT (1992)	CM200-Cs (1998)	Titan 80-300 (2005)	PICO (2012)
300 kV FEG	200 kV FEG	300 kV FEG	300 kV X-FEG
no corrector	Cs corrector	Cs corrector	Cc & Cs corrector

Fig. 9: *Simulation of an AlN crystal viewed along the [113] zone axis using different transmission electron microscopes operated at optimised conditions (Scherzer defocus and Lentzen focus).*

3 Aberration correction in high-resolution STEM

Figure 10 shows a schematic illustration of a scanning TEM (STEM), in which a focused electron beam is scanned across the specimen. For each position of the probe, an on-axis bright field detector or a high-angle annular dark field (HAADF) detector collects electrons that interacted with the specimen. The same microscope can be used for HRTEM and STEM if a microscope such as that shown in Fig. 1 is equipped with a scanning unit located above the specimen and detectors in the diffraction plane. In order to achieve a resolution in the sub-Å range, a STEM has to be equipped with an aberration corrector for the condenser lens.

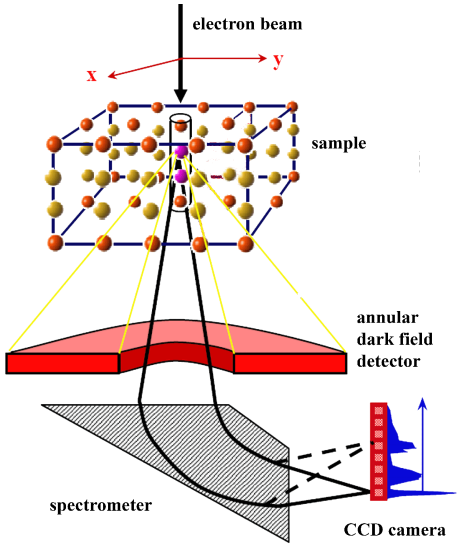


Fig. 10: *In STEM the probe is scanned across the specimen. For each position of the probe, electrons are collected by a bright field or annular dark field detector. A spectrometer can also be used to record an electron energy-loss spectrum.*

The resolution in STEM is determined primarily by the size of the probe. The probe function is in turn determined by parameters that include the diameter of the condenser aperture and the aberrations of the condenser lens system. Correction of the aberrations of the condenser lenses is crucial to obtain sub-Å resolution in HAADF imaging.

Under certain assumptions, the image intensity in annular dark field images is given by the convolution of the square of the probe function and the square of the specimen transmission function. The image formation process can be regarded as incoherent if the inner angle of the annular detector is much larger than the diameter of the condenser aperture, which defines the convergence angle of the illumination. Typically, only electrons that have been scattered to sufficiently large angles on the order of 100 mrad (compared to about 20 mrad for the semi-angle of the condenser aperture) must be collected by the detector. For a sufficiently thin specimen, the image intensity is to a first approximation proportional to the atomic number Z raised to the power 1.7. However, the image can also be affected by beam broadening as the electron probe travels through the specimen and by dynamical diffraction.

A recent application of STEM to the characterization of a complex metallic alloys (CMA) is illustrated in Fig. 11. Such materials have unusual electronic transport [17], magnetic [18] and plastic [19] properties and novel types of dislocations, which are referred to as metadislocations [20]. In the T-Al-Mn-Pd phase, which is orthorhombic with lattice parameters a = 1.47 nm, b = 1.25 nm, and c = 1.26 nm, a deformation mechanism based on the movement of a novel type of dislocation was found. Figure 12 shows an aberration-corrected HAADF STEM image of a dislocation taken along the b direction of the T-phase structure [21, 22]. The unit cell (left blue rectangle) contains 156 atoms. The unit cell can be divided into structural subunits represented by elongated hexagonal tiles (white and yellow polygons) arranged in rows of alternating orientation, mutually tilted by 36°. White dots in the centres of the hexagons correspond to columns containing Pd, while the columns at the edges and vertices of the hexagons contain mainly Mn. The core of the dislocation is represented by a green polygon, which has a complex (1 0 0) mirror symmetric shape and a Burgers vector $\boldsymbol{b} = c/\tau^4 \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ with $|\boldsymbol{b}| = 0.184$ nm, where $\tau = 1/2\left(\sqrt{5}+1\right)$ is the golden mean.

The Burgers vector is a small fraction of the c lattice constant of the T-phase. The dislocation trails a planar defect, which, upon closer inspection at high-magnification, appears to be a nm-thick slab of the so-called R-phase. This R-phase slab is visible at the right side of the dislocation core. On the left side of the dislocation, three bowtie-shaped tiles, shown in red, are required to complete the area-filling tiling. The bowtie-shaped tiles are line defects along the b direction but are not dislocations, since they have no Burgers vectors. During plastic deformation, the bow-tie shaped tiles move ahead of the dislocation core to rearrange the structure, i.e., to clear the way for the dislocation. The dislocation creates a slab of R-phase in its wake and creates a shear-induced transformation from the T- to the R-phase in a slab of about 5 nm thickness. For a single elementary glide step of the dislocation, a coordinated movement of a few hundred atoms is required.

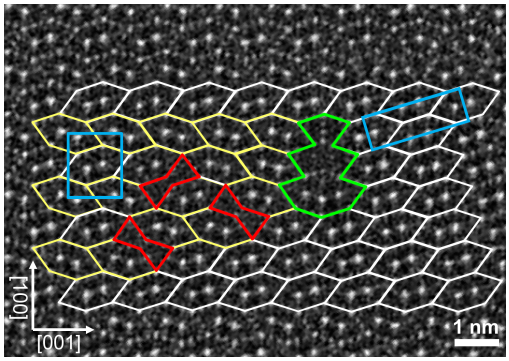


Fig. 11: *High-resolution HAADF-STEM micrograph of a dislocation in CMA T-Al-Mn-Pd. White and yellow hexagons represent structural subunits of the T-phase (left) and R-phase (right). The unit cells are shown by blue rectangles. The dislocation core is represented by a green polygon. The dislocation is escorted by three bowtie-shaped tiles (red).*

A second illustration of the application of aberration-corrected HRTEM and STEM is shown in Fig. 12. The images were obtained from semiconductors that contain transition metal atoms introduced with the intention of combining ferromagnetic and semiconducting properties for the design of spin-electronic devices [23]. In such materials, ferromagnetism can result from the presence of nanoscale clusters of magnetic atoms or randomly located diluted transition metal impurities or defects. In these studies, care was devoted to cross-sectional TEM specimen preparation to minimize artefacts. Figure 12(a) shows an aberration-corrected HRTEM image of a Mn-doped GaAs layer grown by molecular beam epitaxy containing hexagonal (NiAs-type) MnAs crystals, adjacent voids and As nanocrystals. Figure 13(b) shows an aberration-corrected HAADF STEM image of an Fe-doped GaN layer grown by metalorganic chemical vapour deposition containing Fe-rich nitride nanocrystals associated with N_2 filled bubbles.

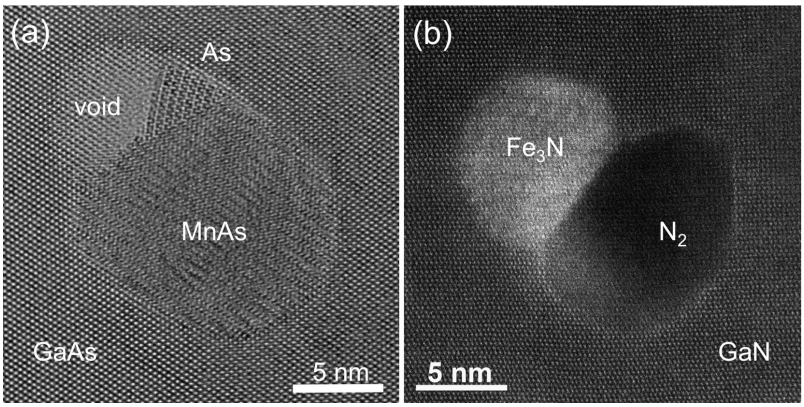


Fig. 12: *(a) Aberration-corrected HRTEM image of hexagonal MnAs, orthorhombic As and a void in a GaAs host. (b) Aberration-corrected HAADF STEM image of an Fe_3N nanocrystal and an adjacent N_2 bubble in a GaN host. The ADF inner semi-angle used was 47.4 mrad.*

4 Electron energy-loss spectroscopy

Electron energy-loss (EEL) spectra can be recorded using either post-column or in-column spectrometers to reveal the inelastic interactions of electrons with the specimen, which involve both energy and momentum transfer. Figure 13 shows the characteristic features in an EEL spectrum. The intensity is plotted as a function of energy loss, with zero energy corresponding to elastically scattered electrons. Small energy losses correspond to valence losses, i.e., plasmon excitations, inter- and intraband transitions. Losses at higher energies are associated with inner shell ionizations. Since the latter signals are very weak compared to the low loss region, different spectrometer settings are used to record the low loss and high loss regions of the spectrum. As the total signal in an ionization edge is proportional to the number of excited atoms, it can be used for quantification of the local chemical composition. Evaluation of the integrated signal requires subtraction of the background in the spectrum. In most cases, a function of the type

$$I(\Delta E, A, r) = A\exp(-r\Delta E) \quad . \tag{8}$$

is used, where A and r are fitting parameters. Details about the quantitative fitting procedures of EEL spectra using calculations of inelastic scattering cross sections and the treatment of specimen thickness effects can be found elsewhere [24].

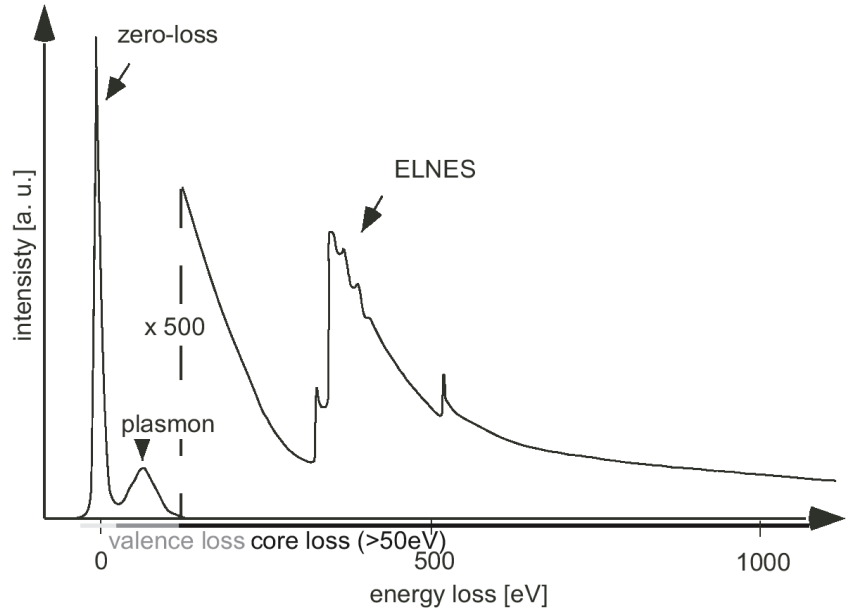


Fig. 13: *Schematic EELS spectrum showing a wide range of energy losses including elastically scattered electrons (zero-loss peak), valence losses (plasmons, inter- and intraband transitions) and the core ionization edges. EELS spectra have a high dynamic range. Note that the ionization edges are superimposed on a background signal.*

The energy loss near edge structure (ELNES) shown in Fig. 13 contains information about the local chemical environment of an excited atom. Two models, displayed Fig. 14, are used to understand edge structure: (a) a band structure model, in which an electron from a shell is excited into unoccupied states above the Fermi level; (b) a multiple scattering model describing ELNES by an outgoing wave, which is scattered by the surrounding atoms. The intensity of the edge scales primarily with the density of final states.

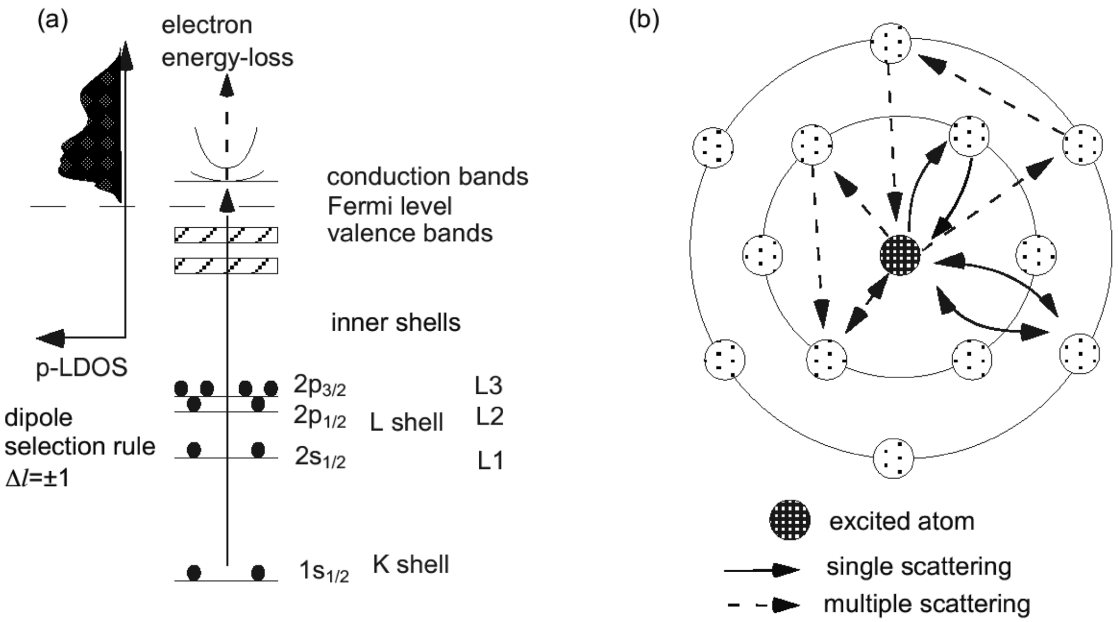


Fig. 14: *Schematic representation of two models for the origin of electron energy-loss near-edge structures (ELNES) for core-ionization edges. (a) Transition of strongly bound core electrons into unoccupied states (b) multiple scattering description of ELNES.*

Figure 15 shows representative background-subtracted EEL spectra recorded from an Fe-doped GaN semiconductor. In this material, depending on the growth temperature used, Fe can either be distributed homogenously in the GaN host lattice or it can accumulate in the form of Fe-N nanocrystals. In the present specimen, Fe-N nanocrystal formation was observed in samples that had been deposited at temperatures higher than 850°C. Most of the Fe-N nanocrystals were found to be associated with closely adjacent void-like features. EELS was used to show that these features are bubbles filled with molecular N_2. In order to interpret the experimental results, N K edge spectra were calculated for GaN using self-consistent real-space multiple-scattering calculations with FEFF 9.05. Figure 15 shows an EELS measurement recorded from a single nanocrystal embedded in the GaN host. A 100 kV accelerating voltage and a distributed-dose acquisition routine [25] were used to minimize and control electron beam induced damage during the experiment. Figure 15(a) shows N K edge spectra recorded from the nanocrystal, the adjacent N_2-containing region and the GaN host. The N K edge shows a three-peaked structure between 400 and 407 eV. Figure 15 (b) shows an experimental spectrum acquired from the bubble alongside an experimental spectrum from N_2 taken from the EELS Atlas [26] and a simulation of a GaN spectrum.

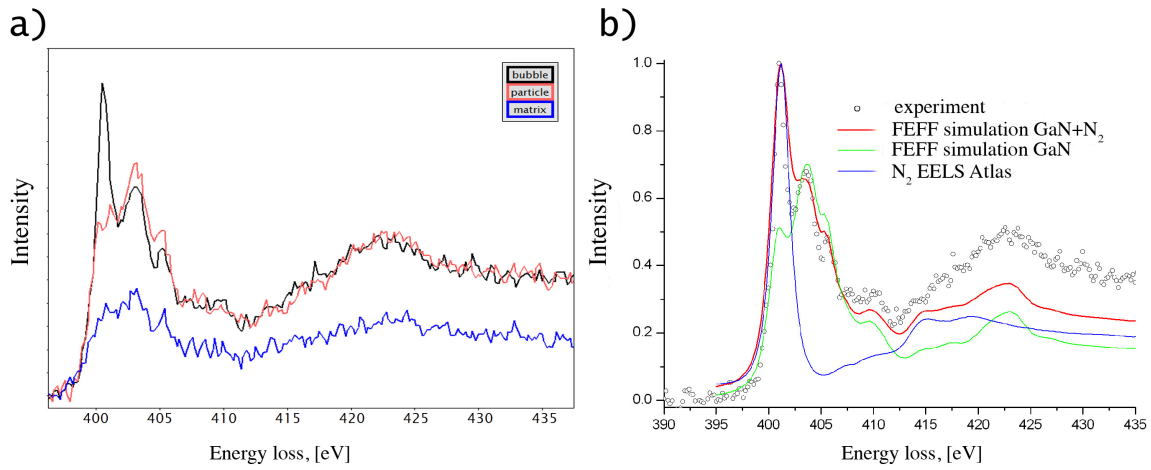


Fig. 15: *(a) Background-subtracted N K edge spectra measured from an Fe-N nanocrystal, a N_2 bubble and the GaN host lattice in Fe-doped GaN. (b) Experimental spectra acquired from a nitrogen bubble alongside an experimental measurement from N_2 taken from the EELS Atlas and simulated spectra calculated for GaN.*

The simultaneous application of high-resolution STEM and EELS can be used to access chemical information about a specimen on the atomic scale. A focused electron beam is scanned across the sample. For each (x,y) position both the HAADF signal and the EEL signal are recorded.

Figures 16-18 illustrate the application of HAADF STEM and EELS to measure the compositions across $DyScO_3/SrTiO_3$ interfaces. The atomic structure of the $DyScO_3$ layer in $SrTiO_3$ is revealed in the aberration corrected HAADF STEM image shown in Fig. 16, which was recorded using a probe size of 0.08 nm. The intensity of each atom column reflects differences in atomic number. Here $DyScO_3$ is imaged along the [101] direction of the orthorhombic unit cell, resulting in a zigzag arrangement of the Dy columns parallel to the interface [27]. Sc and Ti atoms are located at the centre of the oxygen octahedra (displayed in orange in the structure models) and produce weaker contrast compared to Sr and Dy. The oxygen atoms cannot be resolved.

Along the interfaces, in rows 4 and 14, the contrast at the Dy/Sr positions alternates between neighbouring columns. This is seen in the image as well as in the concentration profile. The concentrations were determined by quantification of the intensities using Gaussian fits to the contrast at each atom position in Fig. 16. In each row the intensities of five equivalent positions (A and B) were averaged. Error bars denote the standard deviation obtained from concentration values in regions of constant composition. The Dy concentration was obtained by calculating the integrated intensity to the power of 0.5 and normalizing to 100% within the $DyScO_3$ layer and 0% within the $SrTiO_3$ layer. In row 4 of Fig. 16, the positions labelled “A” are brighter than positions B. The opposite behaviour is seen in row 14, where positions B show the highest intensity, suggesting the presence of an ordered interface structure.

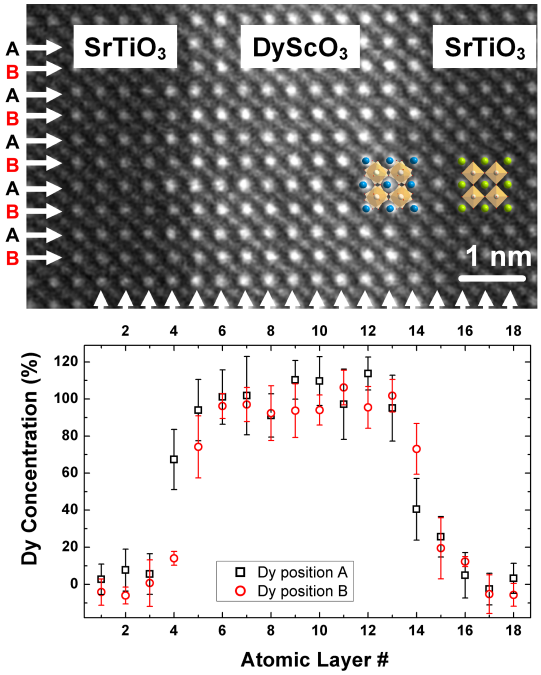


Fig. 16: *HAADF image of a $SrTiO_3/DyScO_3$ multilayer (top). For each Dy layer the concentrations at positions A and B were determined (bottom). The interface layers, rows 4 and 14, show a difference in contrast between neighbouring positions A and B, suggesting an ordered interface structure [28].*

Since Sc and Ti are very close in atomic number (21 and 22, respectively), EEL spectra were recorded simultaneously with the HAADF image in order to assign a spectrum to each atomic layer [29]. Since the edges show only a weak signal, a slightly larger probe size of about 0.2 nm and an acquisition times of 1 s per spectrum were used. Figure 17 shows an HAADF STEM image of $SrTiO_3/DyScO_3/SrTiO_3$. Figure 18 shows EEL spectra recorded from the layers marked in colour in Fig. 17. The spectra show the Sc and Ti L_{23} edges, which have their onsets at 402 eV and 456 eV, respectively. For both elements, four characteristic lines are present, as a result of a splitting into e_g and t_{2g} levels of the L_2 and L_3 edges due to the octahedral crystal field imposed by the O atoms. The Sc signal decreases on moving from the $DyScO_3$ layer into the $SrTiO_3$. The opposite behaviour is seen for the Ti edge. The concentrations of Ti and Sc were determined by extracting the integrated signals of the edges and introducing a normalization, that within the $DyScO_3$ the Sc concentration is 100% and falls to 0% within the $SrTiO_3$. The Ti and Sc concentrations were determined this way for each atomic layer and are plotted in Fig. 19, superimposed on the HAADF image of Fig. 17. The Dy concentration of Fig. 16 is also displayed as an average value for positions A and B in each atomic layer. Neither the Dy nor the Sc concentration changes abruptly at the interface. Instead intermixing is observed, extending over 2-3 atomic layers. Both Sc and Ti show typical edge shapes for octahedral coordination. The fact that the valence of Ti is close to 4+ for all atomic layers can be concluded from the fact that the shapes of the lines do not change. For a Ti^{3+} ion, the line shape would be expected to consist of two broad peaks, instead of the four lines observed. The measured compositions were used to infer the charge distributions in

the layers, and the fact that a predicted ”polar catastrophe” is overcome by an adjustment to the chemical compositions of the individual atomic layers.

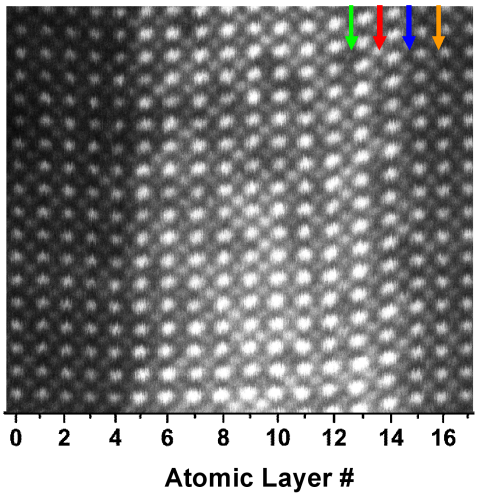


Fig. 17: *Z contrast image of the Dy ScO_3/$SrTiO_3$ multilayer system recorded within 100 s the arrows mark the layers, which correspond to the EEL spectra in Fig 17*

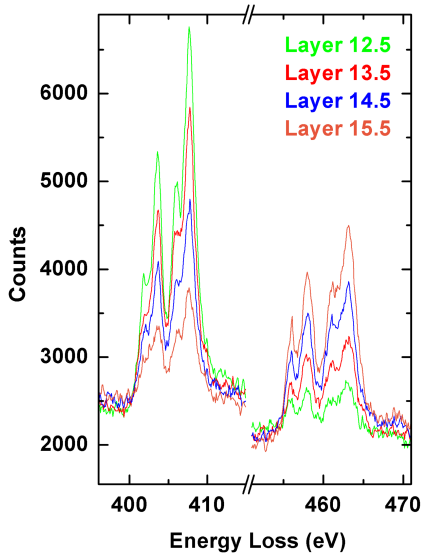


Fig. 18: *EEL spectra of layers 12.5 through 15.5, revealing the Sc L_{23} edge at 402 eV and the Ti L_{23} edges at 456 eV.*

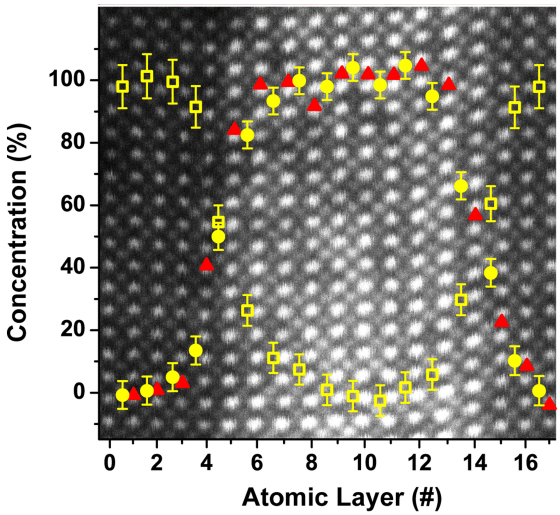


Fig. 19: *HAADF image of $SrTiO_3$/$DyScO_3$. The fast scan direction runs from bottom to top. Superimposed are the average Dy concentrations as red triangles (average of positions A and B in Fig. 16) as well as Sc (yellow circles) and Ti (yellow open squares) concentrations for each atom row. Error bars denote the standard deviation obtained from concentration values in regions of constant composition.*

5 Electron tomography

A significant development in electron tomography for the measurement of the three-dimensional morphology, density or composition of an object is illustrated in Figs 20 and 21. The advance involves the development of an improved alignment procedure for a tilt series of recorded images. Previously, algorithms for three-dimensional reconstruction, such as backprojection or algebraic reconstruction, assumed a precise knowledge of the projection geometry and the absence of any image displacements prior to reconstruction. The most widely used techniques for image alignment in electron tomography were developed for the biological sciences and are based on pattern matching by correlation techniques [30, 31] or by the tracking of fiducial markers [32, 33]. These methods are useful at low and medium magnification, with automated schemes for alignment facilitating the refinement process [34, 35]. However, cross-correlation and marker based methods have disadvantages. Successive cross-correlation of images in a tilt series can only produce an approximate alignment since three-dimensional rotation around a common axis is not taken into account and the image changes with viewing direction. Alignment errors can accumulate during the procedure [36, 37], resulting in a loss of high-resolution detail in the tomogram. Tracking of marker points in sinogram space does account for the three-dimensional rotation but it relies on the presence of object details which can serve as a markers or landmarks. Methods that use the reconstruction feedback of the tomogram iteratively to improve image alignment have been proposed and demonstrated in the literature. These methods include manual feedback control, iterative optimizations in the form of alignment to pseudo-projections of a tomogram sequence [38], three-dimensional model based approaches that align with respect to expected projections of automatically or pre-identified specimen features and area matching as an extension of the cosine stretching used in cross-correlation alignment [39]. A recently developed variant of feedback optimization is the iterative refinement of the tomogram contrast and resolution [40], which does not depend on user assisted motif recognition or the presence of extrinsic markers. Flow charts of the conventional and new algorithm are shown in Fig. 20. The optimization loop of the refinement algorithm provides a correct model of the three-dimensional motion of the object and maximizes the tomogram resolution. Simulation data demonstrate that translational displacement accuracy better than the width of the point spread function in an individual image can be achieved. Experimental tomographic reconstructions at medium resolution show that the marker-free alignment procedure is competitive with fiducial marker alignment and removes artefacts after consecutive cross-correlation alignment [40]. Figure 21 demonstrates the application of the procedure to the reconstruction of WS_2 nanotubes. In this example, no obvious nodal points can be identified and self-markers cannot be utilized. A further complication is that the bundle of nanotubes shows significant motif change in projection during tilt. In such an example, image alignment based on cross correlation is likely to fail. After refinement, the sinogram traces of the tube shells follow the expected projection and no shape distortion is seen in the tomogram slices, apart from residual missing wedge artefacts.

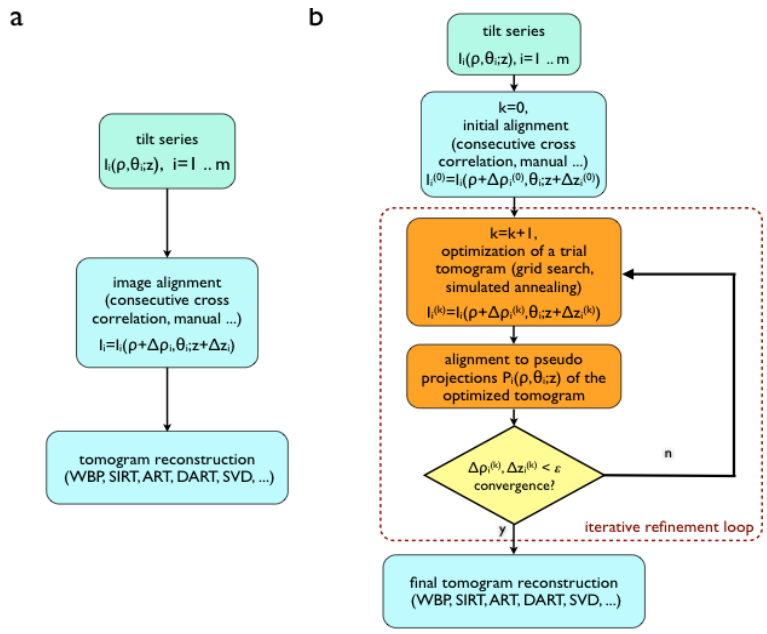


Fig. 20: *Flowchart for refinement of image alignment for tomographic reconstruction from tilt series of images. (a) Conventional procedure involving alignment of images prior to reconstruction. (b) New algorithm including iterative loop.*

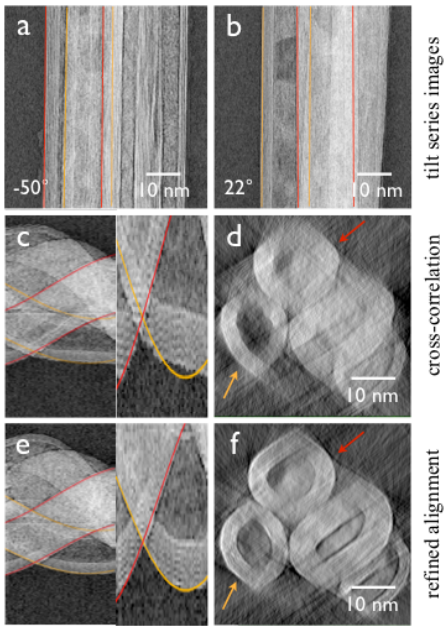


Fig. 21: *Comparison of conventional and new alignment for reconstruction of a bundle of WS_2 nanotubes. (a) and (b) are images from the tilt series. The lines mark the outer shells of two nanotubes. (c) shows a sinogram (left) after conventional cross-correlation alignment. The magnified lower part of the sinogram reveals a lack of correlation. (d) Tomogram slice showing the bundle in cross-section. (e) and (f) Sinogram and a tomogram slice after refinement of the alignment.*

6 Electron holography of magnetic and electric fields

Off-axis electron holography is a specialized TEM technique that can be used to characterize magnetic fields and electrostatic potentials in materials with sub-10-nm spatial resolution [41]. It is the only technique that provides direct access to the phase shift of the electron wave that has passed through a thin specimen, in contrast to TEM techniques that record only spatial distributions of image intensity (described above). The technique can be divided into two primary modes: (i) high-resolution electron holography, in which the interpretable resolution in lattice images is improved by the use of software phase plates to correct for electron microscope lens aberrations [42] and (ii) medium-resolution electron holography, in which magnetic fields and electrostatic potentials within and around materials are studied with nanometre spatial resolution [43]. Some of the most successful early examples of the application of medium-resolution electron holography were the experimental confirmation of magnetic flux quantization in superconducting toroids and the study of magnetic flux vortices in superconductors [44]. Here, more recent developments and applications of medium-resolution electron holography for the study of nanoscale materials are described.

A schematic diagram showing the typical electron-optical configuration for off-axis electron holography is shown in Fig. 22a. The region of interest is positioned so that it covers approximately half the field of view. A biprism (a thin conducting wire), which is usually located close to the first image plane, has a positive voltage applied to it to overlap a "reference" electron wave that passed through vacuum with the electron wave that passed through the specimen. The overlap region contains interference fringes and an image of the specimen, in addition to Fresnel fringes from the edge of the biprism wire (Figs 22b-e). The electron-optical configuration is equivalent to the use of two electron sources S_1 and S_2.

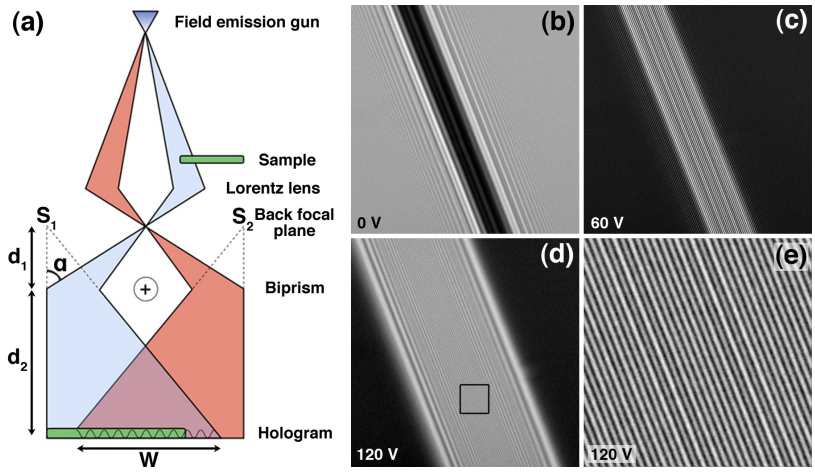


Fig. 22: *(a) Schematic diagram of the setup for off-axis electron holography. (b-e) Interference fringe patterns (electron holograms) recorded from vacuum using biprism voltages of (b) 0 V, (c) 60 V and (d,e) 120 V. In (b), only Fresnel fringes from the edges of the biprism are visible. (e) corresponds to the box marked in (d). The field of view in (b)-(d) is 916 nm.*

The intensity in an off-axis electron hologram consists of the reference image intensity, the specimen image intensity and a set of cosinusoidal fringes. A representative hologram of a specimen that contains a chain of magnetite (Fe_3O_4) crystals is shown in Fig. 23a, alongside a vacuum hologram (Fig. 23b) and a magnified region of the specimen hologram (Fig. 23c). In order to extract phase and amplitude information from an electron hologram, it is Fourier transformed. The Fourier transform contains a peak at the origin corresponding to the Fourier transform of the reference image, a peak at the origin corresponding to the Fourier transform of a bright field image of the specimen, a peak at $\mathbf{q} = -\mathbf{q_c}$ corresponding to the Fourier transform of the image wavefunction and a peak at $\mathbf{q} = +\mathbf{q_c}$ corresponding to the Fourier transform of the complex conjugate of the wavefunction. Figure 23d shows the Fourier transform of the hologram shown in Fig. 23a. In order to recover the complex wavefunction, one of the "sidebands" in the Fourier transform is selected (Fig. 23e) and inverse Fourier transformed. The amplitude and phase are then calculated. The phase image is initially calculated modulo 2π, meaning that 2π discontinuities appear where the phase shift exceeds this amount (Fig. 23g). The phase image can be "unwrapped" by using suitable algorithms, as shown in Fig. 23h. In Fig. 23d, the streak from the "centerband" toward the sidebands is attributed to the presence of Fresnel fringes from the biprism wire, which can lead to artefacts in the reconstructed amplitude and phase. This effect can be minimized by masking the streak from the Fourier transform before inverse Fourier transformation (Fig. 23e).

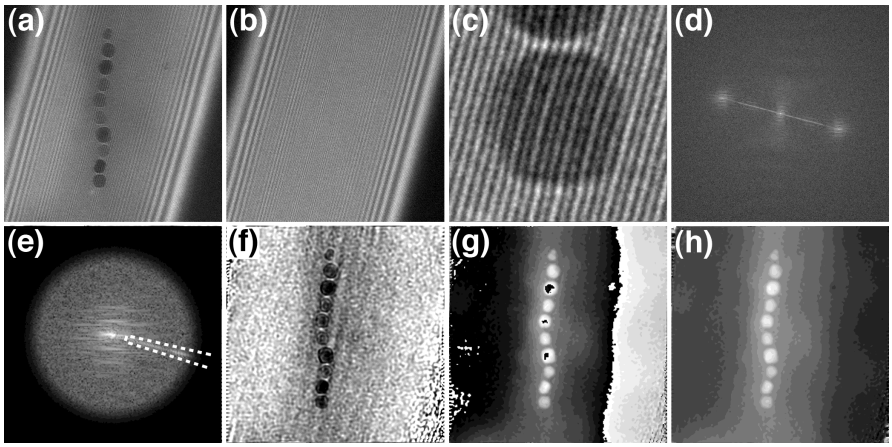


Fig. 23: *Processing steps required to convert a recorded electron hologram into an amplitude and phase image. (a) Hologram of a chain of magnetite crystals. The overlap width and holographic interference fringe spacing are 650 and 3.3 nm, respectively. The field of view is 725 nm. (b) Vacuum hologram acquired after the specimen hologram. (c) Magnified region of the specimen hologram. (d) Fourier transform of (a). (e) One of the sidebands extracted from the Fourier transform, shown after applying a circular mask with smooth edges. The streak from the Fresnel fringes can be removed by assigning a value of zero to pixels inside the region shown by the dashed line. Inverse Fourier transformation of the sideband is used to provide a complex image wave, which is displayed in the form of (f) an amplitude image and (g) a modulo 2π phase image. Phase unwrapping algorithms are used to remove the 2π phase discontinuities from (g) to yield an unwrapped phase image shown in (h).*

The phase shift recorded using electron holography is sensitive to both the electrostatic potential and the in-plane component of the magnetic induction in the specimen integrated in the beam direction. Neglecting dynamical diffraction (i.e., assuming that the specimen is thin and weakly diffracting), the phase shift can be written in the form

$$\phi(x,y) = \phi_e(x,y) + \phi_m(x,y) = C_E \int_{-\infty}^{+\infty} V(x,y,z)\,dz - \frac{e}{\hbar}\int_{-\infty}^{+\infty} A_z(x,y,z)\,dz \quad . \tag{9}$$

where V is the electrostatic potential, A_z is the component of the magnetic vector potential parallel to the electron beam direction,

$$C_E = \left(\frac{2\pi}{\lambda}\right)\left(\frac{E+E_0}{E(E+2E_0)}\right) \quad . \tag{10}$$

is a constant that depends on the accelerating voltage $U=E/e$, λ is the (relativistic) electron wavelength and $E_0 = 511$ keV is the rest mass energy of the electron. C_E takes a value of 6.53×10^6 rad/m^{-1}V^{-1} at an accelerating voltage of 300 kV. If no external charges or applied electric fields are present, then the only electrostatic contribution to the phase originates from the mean inner potential V_0 coupled with variations in the specimen thickness $t(x,y)$. If the specimen has a uniform composition, then the electrostatic contribution to the phase is proportional to $t(x,y)$. If $t(x,y)$ or its gradient is known then φ_e can be used to measure V_0. The magnetic contribution to the phase shift carries information about the magnetic flux. The difference between its value at any two points (x_1,y_1) and (x_2,y_2)

$$\Delta\phi_m = \phi_m(x_1,y_1) - \phi_m(x_2,y_2) = -\frac{e}{\hbar}\int_{-\infty}^{+\infty} A_z(x_1,y_1,z)\,dz + \frac{e}{\hbar}\int_{-\infty}^{+\infty} A_z(x_2,y_2,z)\,dz \quad . \tag{11}$$

can be written in the form

$$\Delta\phi_m = -\frac{e}{\hbar}\oint \mathbf{A}\cdot d\mathbf{l} \quad . \tag{12}$$

for a rectangular loop formed by two parallel electron trajectories crossing the sample at coordinates (x_1,y_1) and (x_2,y_2) and joined, at infinity, by segments perpendicular to the trajectories. By virtue of Stokes' theorem,

$$\Delta\phi_m = \frac{e}{\hbar}\iint \mathbf{B}\cdot\hat{\mathbf{n}}\,dS = \frac{\pi}{\varphi_0}\Phi(S) \quad . \tag{13}$$

where $\varphi_0=h/2e=2.07\times10^{15}$ Tm^2 is a flux quantum. The difference between any two points in a phase image is therefore a measure of the magnetic flux through the region of space bounded by two electron trajectories crossing the sample at the positions of these two points and a graphical representation of the magnetic flux distribution can be obtained by adding contours

to a phase image. A phase difference of 2π corresponds to an enclosed magnetic flux of 4.14×10^{15} Tm^2.

The relationship between the magnetic contribution to the phase shift and the magnetic induction can be established from the gradient of φ_m. In the special case when stray fields can be neglected, the sample has a constant thickness and the magnetic induction does not vary with z in the specimen, the magnetic contribution to the phase gradient is proportional to the magnetic induction.

When characterizing magnetic materials, the microscope objective lens is usually turned off because it creates a large (>2 T) magnetic field at the position of the specimen. Instead, a non-immersion Lorentz lens can be used as the imaging lens. This lens allows specimens to be imaged at relatively high magnification in magnetic-field-free conditions. The spherical aberration coefficient of this lens is large (e.g., 8400 mm for an FEI Titan), resulting in a point resolution of ~2 nm. It should be noted that the separation of electrostatic and magnetic contributions to the phase shift is almost always mandatory in order to obtain quantitative magnetic information from a phase image. The few instances when this step may be avoided include the special case of magnetic domains in a film of constant thickness (far from the specimen edge). The contribution of the mean inner potential to the phase shift can be much greater than the magnetic contribution for small (sub-50-nm) magnetic nanocrystals. Separation of the magnetic contribution from the recorded phase shift is then essential. In principle, the most accurate way of achieving this separation involves turning the specimen over after acquiring a hologram and acquiring a second hologram from the same region. The sum and difference of the holograms are then used to determine twice the mean inner potential and twice the magnetic contribution to the phase shift, respectively. An alternative method, which is often more practical, involves performing a magnetization reversal experiment in the microscope and then selecting pairs of holograms that differ only in the magnetization direction in the specimen. If the magnetization in the specimen does not reverse perfectly, then such reversal measurements may need to be repeated many times so that non-systematic differences between reversed images are averaged out.

Figure 24 shows magnetic contributions to the phase shift and magnetic induction maps measured from three self-assembled rings of 25-nm-diameter polycrystalline Co nanoparticles using off-axis electron holography [45]. Each particle comprises a core of hexagonal-close-packed Co and a shell of ~3 nm of CoO. Holograms were recorded in zero-field conditions after applying chosen out-of-plane magnetic fields *in situ* in the TEM by partially exciting the conventional microscope objective lens. The mean inner potential contribution to the phase shift was calculated from phase images that had been acquired before and after turning the specimen over. Their sum and difference were used to determine twice the mean inner potential and twice the magnetic contribution to the phase shift, respectively. Once the mean inner potential contribution had been determined, it was subtracted from all subsequent phase images acquired from the same region of the specimen. The resulting phase images were smoothed slightly and phase contours and colours were used to form the final induction maps. An out-of-plane field of -20,000 Oe was initially applied perpendicular to the plane of the specimen and reduced to zero. The sample was then taken out of the TEM in zero field,

turned over and put back into the microscope. Chosen out-of-plane fields were then applied to the specimen in the TEM in succession by changing the current in the TEM objective lens. The applied field was reduced to zero before recording each hologram. Magnetization reversal of most rings could be achieved by using out-of-plane fields of between 1600 and 2500 Oe.

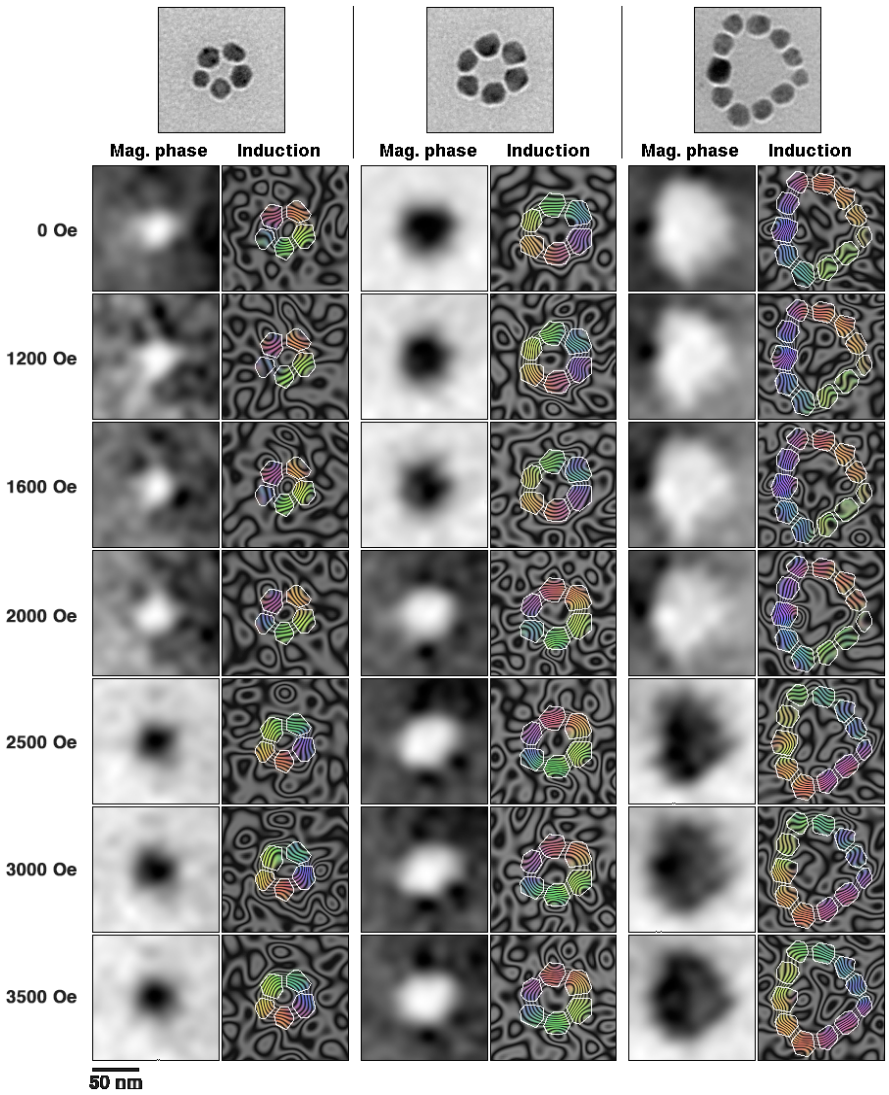


Fig. 24: *Top row: Bright-field TEM images of rings of polycrystalline Co nanoparticles. Subsequent rows: Magnetic contributions to the phase shift and magnetic induction maps recorded with the specimen in magnetic-field-free conditions using off-axis electron holography after saturating the rings using a large (-20,000 Oe) out-of-plane magnetic field and then applying the indicated out-of-plane fields in succession. The direction of the projected magnetic induction is shown using both contours and colours. (Red=right; blue=up; green=left; yellow=down). The magnetic phase contour spacing is 0.065 rad, corresponding to 96 × phase amplification.*

Figure 25 shows the application of electron holography to study the current-induced motion of transverse and vortex-type magnetic domain walls in permalloy (NiFe) zigzag lines on Si_3N_4 windows [46]. Figure 25(b) shows the sequential positions of a magnetic domain wall that was subjected to 10 μs pulses with a current density of 3.14×10^{11} A/m^2. Electron holograms were acquired at each of these positions. A transverse magnetic domain wall initially formed at a kinked region of the wire after the application of a magnetic field. After applying a current pulse, the wall moved by 2330 nm in the direction of electron flow and transformed into a vortex-type wall. After a second pulse, the vortex-type wall moved in the same direction and became distorted, with the long axis of the vortex increasingly perpendicular to the wire length. This behaviour may be associated with edge roughness or defects, which may restrict movement of the wall. After a third pulse, the domain wall moved 260 nm further and retained its vortex state.

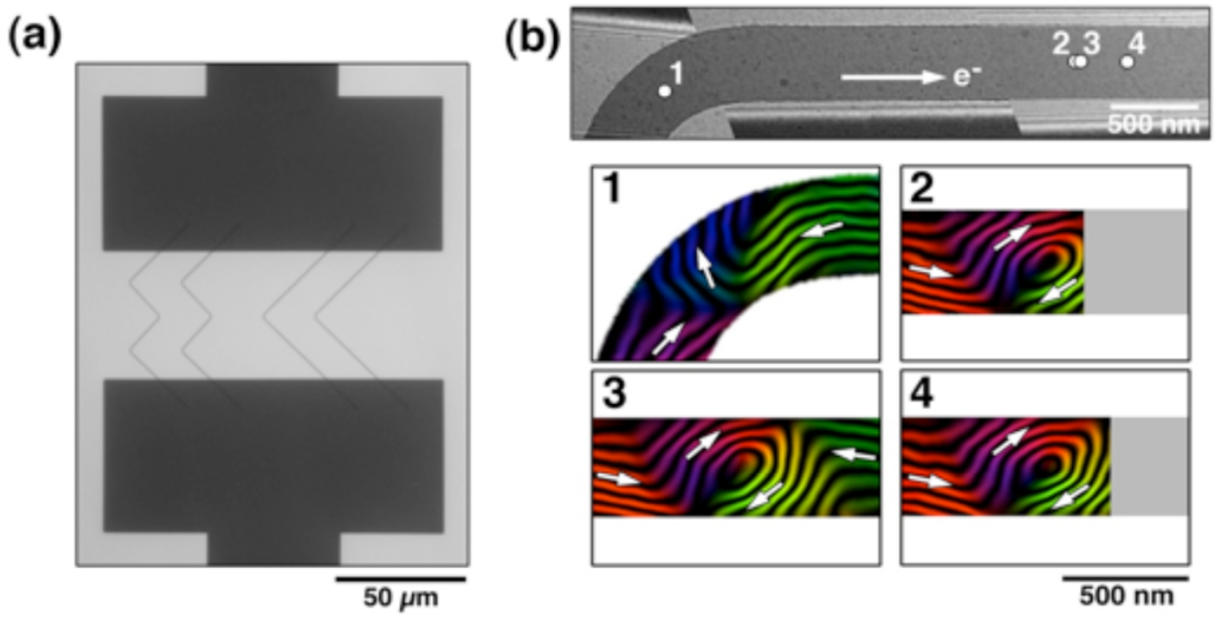


Fig. 25: *(a) Low magnification bright-field image of permalloy (NiFe) zigzag wires (width: 430 nm, thickness: 11 nm) patterned on a Si_3N_4 window using lithography. A kink in one wire was examined using electron holography in the TEM. (b) Higher magnification bright-field image of a permalloy wire and magnetic induction maps measured using off-axis electron holography, showing the positions of a magnetic domain wall after the injection of 10 μs current pulses with a current density of 3.14×10^{11} A/m^2. The arrow in the top image indicates the direction of electron flow. The phase contour spacing in the magnetic induction maps is 0.785 radians.*

One of the most important applications of medium-resolution off-axis electron holography is the measurement of electrostatic potentials in semiconductor devices, in order to fulfil the requirements of the semiconductor industry for a quantitative dopant profiling technique that has nm-scale spatial resolution. Such measurements are often made from site-specific regions

of semiconductor devices by using focused ion beam milling with Ga ions to prepare the TEM specimen. Unfortunately, focused ion beam milling can affect the electrical properties of the specimen as a result of the creation of defects that can extend to a depth of more than 100 nm. Approaches such as low energy ion milling can then be used to reduce the thicknesses of the damaged regions, although they do not remove them entirely. The effects of specimen preparation and illumination by the high energy electron beam on the measured electrical properties are presently being addressed by several research groups [48]. Figure 25 illustrates the recent application of electron holography to the measurement of the electrostatic potential within and around a carbon nanotube that has a voltage applied to it *in situ* in the TEM [49]. Figures 26 (a) and (b) show defocused bright-field images of the specimen, while Fig. 26(c) shows a phase image of one nanotube acquired using off-axis electron holography. The charge distribution along the nanotube was determined by analysing the gradient of the recorded phase image. In this way, the nanotube in Fig. 26(c) was inferred to have two different charge densities along its length (2.3 and 4.5 e/nm).

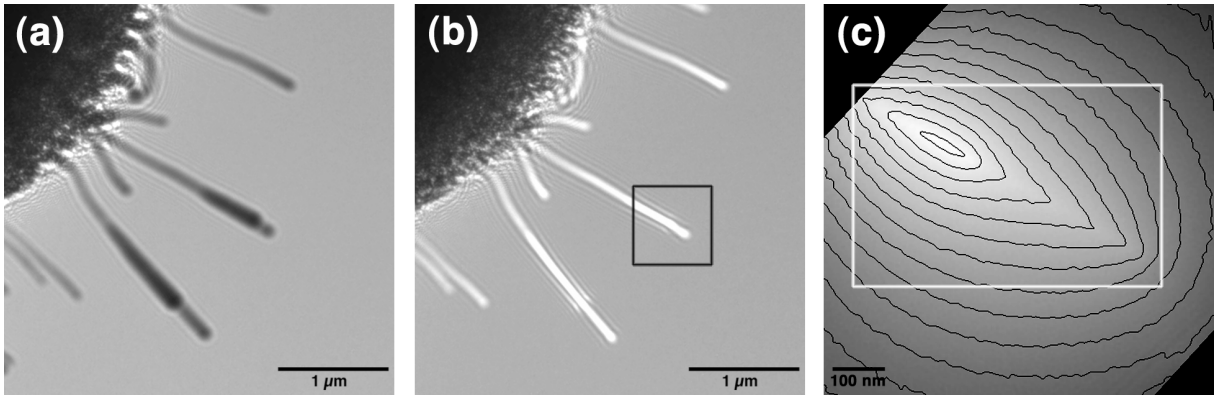


Fig. 26: *(a) Underfocus and (b) overfocus images (defocus ~3 mm) of carbon nanotubes protruding from a nanotube bundle with an applied bias of 50 V created by a Au electrode (not shown) 5 μm away. (c) Phase image acquired from the region indicated in (b) with 1.5 rad phase contours superimposed.*

Figure 27 illustrates the application of the recently developed technique of dark-field electron holography to measure the strain distribution in an InAs quantum dot grown in InP using low pressure metal organic vapour-phase epitaxy at 520 °C. The maximum strain measured in the centre of the dot, relative to the unstrained substrate, is 5.4±0.1%. Comparisons of the strain measurements with computer simulations suggest that the As concentration in the wetting layer is 29% InAs and the composition of the centre of the dot is 100% InAs [50].

Recent developments in medium resolution off-axis electron holography include the development of new approaches for the acquisition, analysis and simulation of electron holograms, the design and use of specimen holders that allow electrical contacts to be applied to magnetic devices *in situ* in the electron microscope, and the comparison of recorded magnetic induction maps both with micromagnetic simulations and with three-dimensional

information about the local compositions and morphologies of the same specimens acquired using electron tomography. In the future, the development electron holography will require new approaches for the careful separation of weak magnetic signals from unwanted contributions to recorded phase images. Further work is also required to increase the sensitivity of the technique for measuring weak fields and to improve its time resolution. The prospect of characterizing magnetic vector fields *inside* nanocrystals in three dimensions by combining electron tomography with electron holography is also of great interest.

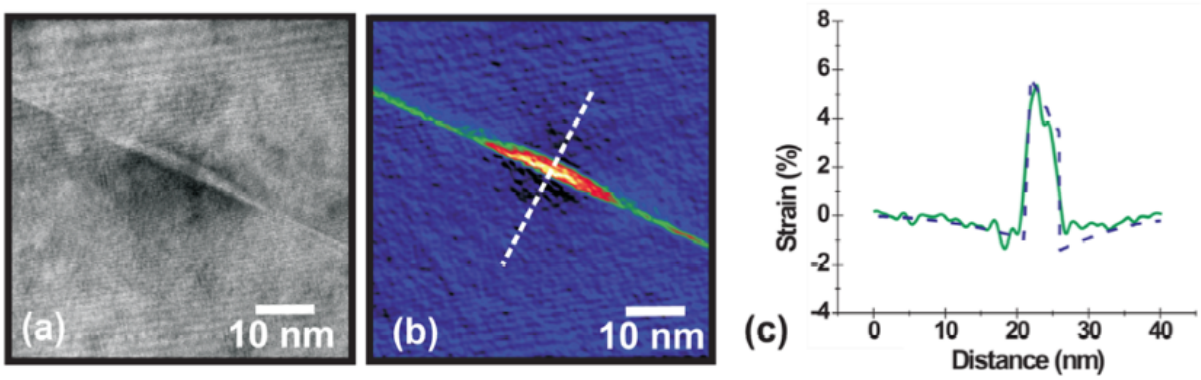


Fig. 27: *(a) Dark-field off-axis electron hologram of an InAs quantum dot in InP, with the 004 diffraction spot selected. (b) Strain map for the 004 growth direction derived from (a). (c) Strain profile extracted from the region indicated by the dashed line in (b) (solid line) compared to simulation (dashed line).*

Acknowledgments

We are grateful to J. Barthel, M. Bar Sadan, G. Pozzi, A. Wei, P. A. Midgley, D. Cooper , M.R. McCartney. C.B. Boothroyd. B. Schaffer, M. S. Moreno, T. Kasama, M. Beleggia, A.J. Craven, M. Kläui, J.L. Rouvière, A. Béché, S. Kadkhodazadeh, E.S. Semenova, K. Yvind, M. Eltschka, M. Wötzel, L.J. Heyderman, S. Hoffmann, A. Bonanni, T. Dietl and other colleagues for discussions and ongoing collaborations.

References

[1] J.C.H. Spence, High-Resolution Electron Microscopy (Oxford University Press, New York, 2003)
[2] L. Reimer and H. Kohl, Transmission Electron Microscopy: Physics of Image Formation (Springer, New York, 2008)
[3] D.B. Williams and C.B. Carter, Transmission Electron Microscopy: A Textbook for Materials Science (Plenum Press, New York, 2009)

[4] A. Thust, in: Vorlesungsmanuskripte des 29. Ferienkurses des Instituts für Festkörperforschung, Schriften des Forschungszentrums Jülich, Reihe Materie und Material, Band 1, 1998
[5] H. Bethe, Ann. Phys. 87, 55 (1928)
[6] O. Scherzer, Z. Phys. 101, 593 (1936)
[7] O. Scherzer, J. Appl. Phys. 20, 20 (1949)
[8] P. Schiske, Proc. 4th Eur. Conf. on Electron Microscopy, Rome, 145 (1968)
[9] W.M.J. Coene, A. Thust, M. Op de Beeck and D. van Dyck, Ultramicroscopy 64, 109 (1996)
[10] A. Thust, W.M.J. Coene, M. Op de Beeck and D. van Dyck, Ultramicroscopy 64, 211 (1996)
[11] C. Kisielowski, C.J.D. Hetherington, Y.C. Wang, R. Kilaas, M.A. O'Keefe, A. Thust, Ultramicroscopy 89, 243 (2001)
[12] M. Haider, H. Rose, S. Uhlemann, E. Schwan, B. Kabius and K. Urban, Ultramicroscopy 75, 53 (1998)
[13] M. Lentzen, B. Jahnen, C.L. Jia, A. Thust, K. Tillmann and K. Urban, Ultramicroscopy 92, 233 (2002)
[14] C. L. Jia, M. Lentzen, and K. Urban, Science 299, 870 (2003)
[15] C.L. Jia, S.B. Mi, K. Urban, I. Vrejoiu, M. Alexe and D. Hesse, Nature Materials 7, 57 (2008)
[16] J. Barthel and A. Thust, Ultramicroscopy 111, 27 (2010)
[17] M. Andersson, M. Feuerbacher and Ö. Rapp, Phys. Rev. B 78, 024201 (2008)
[18] J. Dolinšek, J. Slanovec, Z. Jagličić, M. Heggen, S. Balanetskyy, M. Feuerbacher and K. Urban, Phys. Rev. B 77, 064430 (2008)
[19] S. Roitsch, M. Heggen, M. Lipinska-Chwalek and M. Feuerbacher, Intermetallics 15, 833 (2007)
[20] M. Feuerbacher and M. Heggen, Phil. Mag. 86, 935 (2006)
[21] M. Heggen, L. Houben and M. Feuerbacher, Nature Materials 9, 332 (2010)
[22] M. Heggen, L. Houben and M. Feuerbacher, Acta Mater. 59, 4458 (2011)
[23] T. Dietl, Nature Materials. 9, 965 (2011)
[24] R.F. Egerton, Electron Energy-Loss Spectroscopy in the Electron Microscope (Springer, New York, 2011)
[25] K. Sader, B. Schaffer, G. Vaughan, R. Brydson, A. Brown and A. Bleloch, Ultramicroscopy 110, 998 (2010)
[26] C.C. Ahn and O.L. Krivanek, EELS Atlas (Gatan, 1983)
[27] M. Boese, T. Heeg, J. Schubert and M. Luysberg, Journal of Materials Science 41, 4434 (2006)
[28] M. Luysberg, M. Heidelmann, L. Houben, M. Boese, T. Heeg, J. Schubert and M. Roeckerath, Acta Materialia 57, 3192 (2009)
[29] M. Heidelmann, J. Barthel and L. Houben, Ultramicroscopy 109, 1447 (2009)
[30] R. Guckenberger, Ultramicroscopy 9, 167 (1982)
[31] J. Frank and B.F. McEwan, in: J. Frank (Ed.), Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope (Plenum Press, New York, 1992)

[32] M.C. Lawrence, in: J. Frank (Ed.), Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope (Plenum Press, New York, 1992)
[33] D.N. Mastronarde, in: J. Frank (Ed.), Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell (Springer: New York, 2006)
[34] J.C. Fung, W. Liu, W.J. De Ruijter, H. Chen, C.K. Abbey, J.W. Sedat and D.A. Agard, Journal of Structural Biology 116, 181 (1996)
[35] S. Brandt and U. Ziese, Journal of Microscopy 222, 1(2006)
[36] W.O. Saxton, W. Baumeister and M. Hahn, Ultramicroscopy 13, 57 (1984)
[37] J. Frank, B.F. McEwen and M. Radermacher, Journal of Electron Microscopy Technique 6, 193 (1987)
[38] C.H. Owen and W.J. Landis, Ultramicroscopy 63, 27 (1996)
[39] H. Winkler and K.A. Taylor, Ultramicroscopy 106, 240 (2006)
[40] L. Houben and M. Bar Sadan, Ultramicroscopy 111, 1512 (2011)
[41] D. Gabor, D, Proc. Roy. Soc. A, 197, 454 1949)
[42] M. Lehmann, H. Lichte, D. Geiger, G. Lang and E. Schweda, Mater. Character., 42, 249 (1999)
[43] R.E. Dunin-Borkowski, M.R. McCartney and D.J. Smith, in H.S. Nalwa (Ed.), Encyclopedia of Nanoscience and Nanotechnology, Vol. 3, pp.41-100 (American Scientific Publishers, Stevenson Ranch, 2004).
[44] T. Matsuda, S. Hasegawa, M. Igarashi, T. Kobayashi, M. Naito, H. Kajiyama, J. Endo, N. Osakabe, A. Tonomura and R. Aoki, Phys. Rev. Lett. 62, 2519 (1989)
[45] T. Kasama, R.E. Dunin-Borkowski, M.R. Scheinfein, S.L. Tripp, J. Liu and A. Wei, Adv. Mater. 20, 4248 (2008)
[46] F. Junginger, M. Kläui, D. Backes, U. Rüdiger, T. Kasama, R.E. Dunin-Borkowski, L.J. Heyderman, C.A.F. Vaz and J.A.C. Bland, Appl. Phys. Lett. 88, 212510 (2007)
[47] M. Eltschka, M. Wötzel, J. Rhensius, S. Krzyk, U. Nowak, M. Kläui, T. Kasama, R.E. Dunin-Borkowski, L.J. Heyderman, H.J. van Driel, and R.A. Duine, Phys. Rev. Lett. 105, 056601 (2010)
[48] D. Cooper, R. Truche, P. Rivallin, J.M. Hartmann, F. Laugier, F. Bertin, A. Chabli and J.L. Rouviere, Appl. Phys. Lett. 91, 143501 (2007)
[49] M. Beleggia, T. Kasama, R.E. Dunin-Borkowski, S. Hofmann and G. Pozzi, Appl. Phys. Lett. 98, 243101 (2011)
[50] D. Cooper, J.L. Rouvière, A. Béché, S. Kadkhodazadeh, E.S. Semenova, K. Yvind and R.E. Dunin-Borkowski, Appl. Phys. Lett. 99, 261911 (2011)

F 2 Imaging with X-rays, Neutrons and Synchrotron Radiation

E. H. Lehmann
Neutron Imaging & Activation Group
Spallation Neutron Source Division
Paul Scherrer Institut, Switzerland

Contents

Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

1 Introduction

More than one century ago, Konrad Röntgen discovered X-rays and their performance – by means of imaging methods. His first pictures of human bones inside the body initiated a revolution in medical inspections. It is still common in all hospitals on very routine basis to perform transmission images of parts of the patient's body to find abnormalities and defects.
The use of X-rays for material research on the microscopic and atomic scale was initiated by Laue, Bragg and Debye more than ten years later. Even than, X-ray films were used initially for the investigation of the diffraction rings and spots around the crystalline samples.
The discovery of neutrons (Chadwick) came about 40 years after Röntgen's experiments and found most interests in atomic physics, reactor physics, nuclear weapon programs and radiochemistry. The hype for nuclear weapons at the end of the Second World War and the installation of reactors provided also the access to beams of free neutrons. However, first neutron images were done in Berlin by Kalman and Kuhn at an accelerator driven neutron source.
Because the radiography with X-rays and neutrons were done until about 1970 exclusively with film methods, the applications were limited in respect to the frame rate, the image quality and the introduction of more advanced methods.

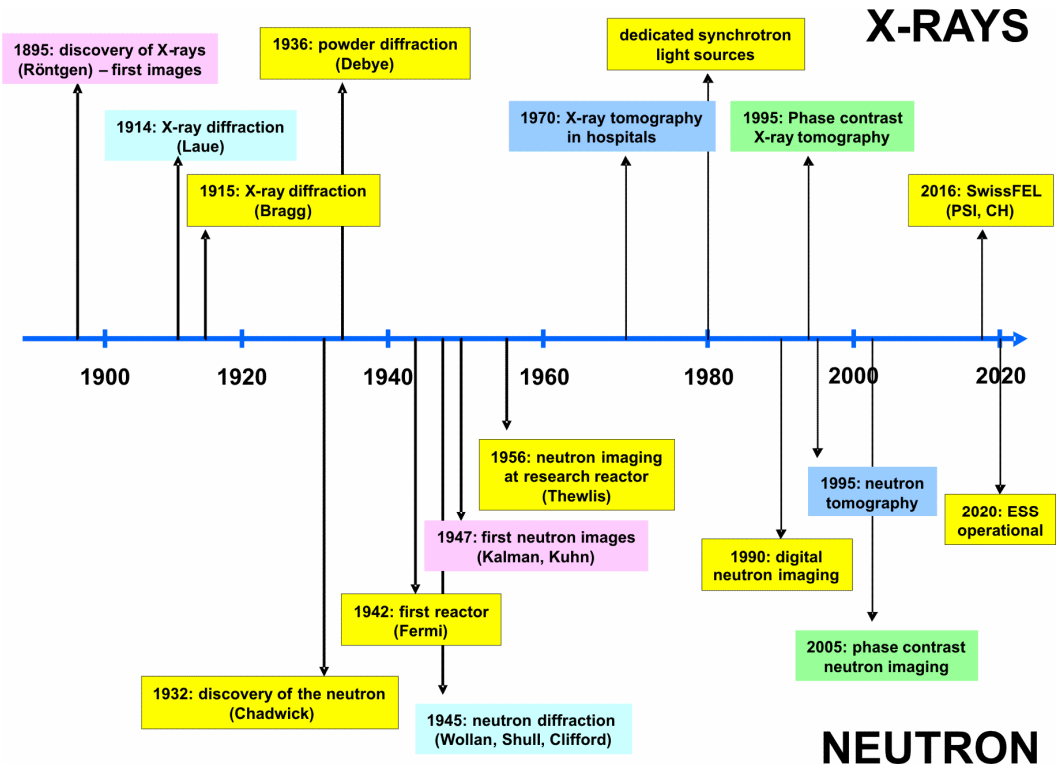


Fig.1: *Historical overview in respect to X-ray and neutron imaging aspects (selection of facts)*

Diffraction and other kinds of scattering methods with neutrons were introduced into material research according to the availability of suitable sources. Neutron scattering has become an established research field by using reactor sources like at ILL (Grenoble) or FRM-2 (Munich) and at other prominent places around the world. About 250 reactor and accelerator based neutron sources are counted by the IAEA [1] in 2010.

X-ray sources on the basis of tubes are limited in intensity and flexibility. The invention and installation of electron accelerators for the production of synchrotron light have been a huge step forward for the X-ray diffraction and imaging community in order to increase the performance and enabled many new techniques. This development will be topped at some places by means of free electron laser facilities, preferentially working in the time domain.

On the other hand, it was another revolutionary step forward to enable the detection of the radiation with digital systems of much higher performance than films. In particular, the imaging techniques for both neutrons and X-rays took profit from this development and can provide today a large variety of new methods with high potential in many applications.

Fig. 1 summaries some most important inventions and activities in respect to X-ray and neutron research since the discovery of these radiations.

This article intends to describe some basics in respect to modern imaging techniques using X-rays, neutrons and synchrotron radiation. Based on the principle of interaction of matter with the different kinds of radiation it will be explained which information can be obtained by the inspection of material samples.
Technical details about facilities, imaging techniques and current applications will be given on the basis of the author's experiences at the installations at Paul Scherrer Institut (Switzerland).
The outlook is focussed on new trends in the methodical developments, the installation of new imaging systems and the creation of new application fields.

2 Basic principles

All transmission radiation techniques (with neutrons, X-rays and synchrotron radiation) are based on the same principle. A specially tuned beam of the radiation is selected from the source by means of collimator devices and filters. After the interaction with the object of interest the transmitted radiation is detected with a two-dimensional area detector, sensitive for the specific kind of radiation. This procedure is schematically described in Fig. 2.

Two beam geometries are interesting for practical applications: a parallel beam which allows a 1:1 shadow image of the object without geometrical blurring and a cone beam with a narrow focussing source point which enables a magnified image of the object. In reality, both approaches can only be made available with compromises due to the limitation of the radiation sources (source size and intensity). A collimation ratio CR can be defined as:

$$CR = \frac{L}{D} \tag{1}$$

D corresponds to the size of the primary aperture and L to the length of collimation between the aperture and the detector. This number should be larger than 100 in order to obtain reasonable spatial resolution. The geometric blurring u is related to the sample-detector-distance d like:

$$u = \frac{d}{CR} \tag{2}$$

The obtained field-of-view (FOV) is given by the beam size and should be at least as large as the detector area. It depends on the object size and its composition which detector area is most suitable for the investigation with the particular kind of radiation – see below.

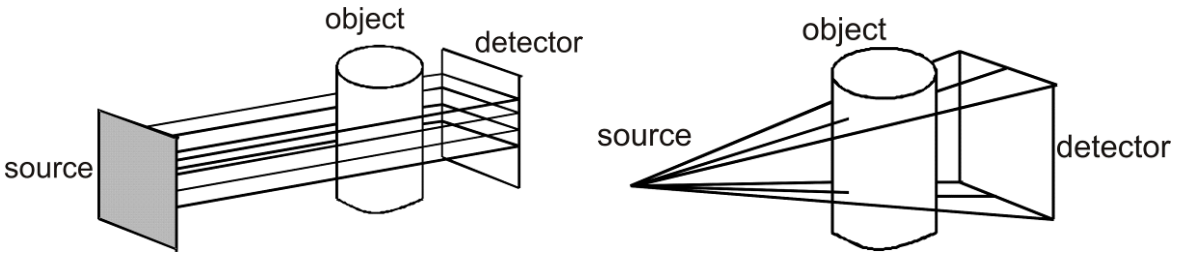


Fig. 2: *The principles in transmission radiation imaging: parallel beam (left); cone beam (right)*

For a quantitative approach in the transmission imaging techniques, the exponential attenuation law according to Beer - Lambert has to be taken into account:

$$I(x,y,E) = I_0(x,y,E) \cdot e^{-\Sigma(x,y,E) \cdot d(x,y)} \tag{3}$$

The intensity of the initial beam I_0 arriving at a position in x and y into the z direction with an energy distribution around E is related to the transmitted beam I according to that exponential law. The exponent is given by the material thickness d in beam direction z and the attenuation coefficient Σ which is a material parameter with a high dependence on the beam energy.

The interaction of X-rays (and synchrotron radiation) with matter takes place in the electron shell of the atoms (mainly Compton scattering and photo-electron emission). Therefore it is understandable that atoms with fewer electrons in their shell interact only weakly with X-rays than these of heavy elements. The principle energy dependency of the attenuation for X-rays in the range between 1 keV and 100 MeV is shown with the example of Al in Fig. 3. The logarithmic scale has to be taken into account. Other materials have a similar slope but show different absolute values.

It is common to list the X-ray data for the attenuation coefficients with μ (unit cm^{-1}) or sometimes with μ /ρ (ρ=material density) as mass-attenuation coefficient (unit cm^2/g). In the case of neutron interactions Σ is called macroscopic cross-section (unit cm^{-1}) which is related to the tabulated microscopic cross-sections σ as follows:

$$\Sigma = N \cdot \sigma = \frac{\rho \cdot L}{M} \cdot \sigma \qquad (4)$$

N corresponds to the density of nuclei, M is the atomic weight and L the Avogadro's constant.

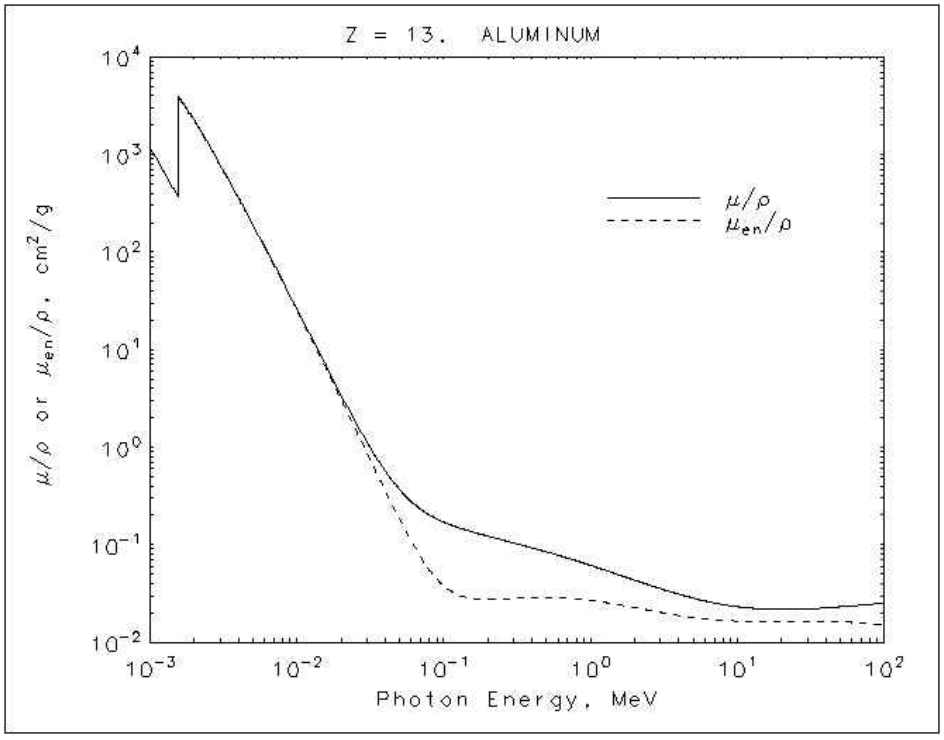
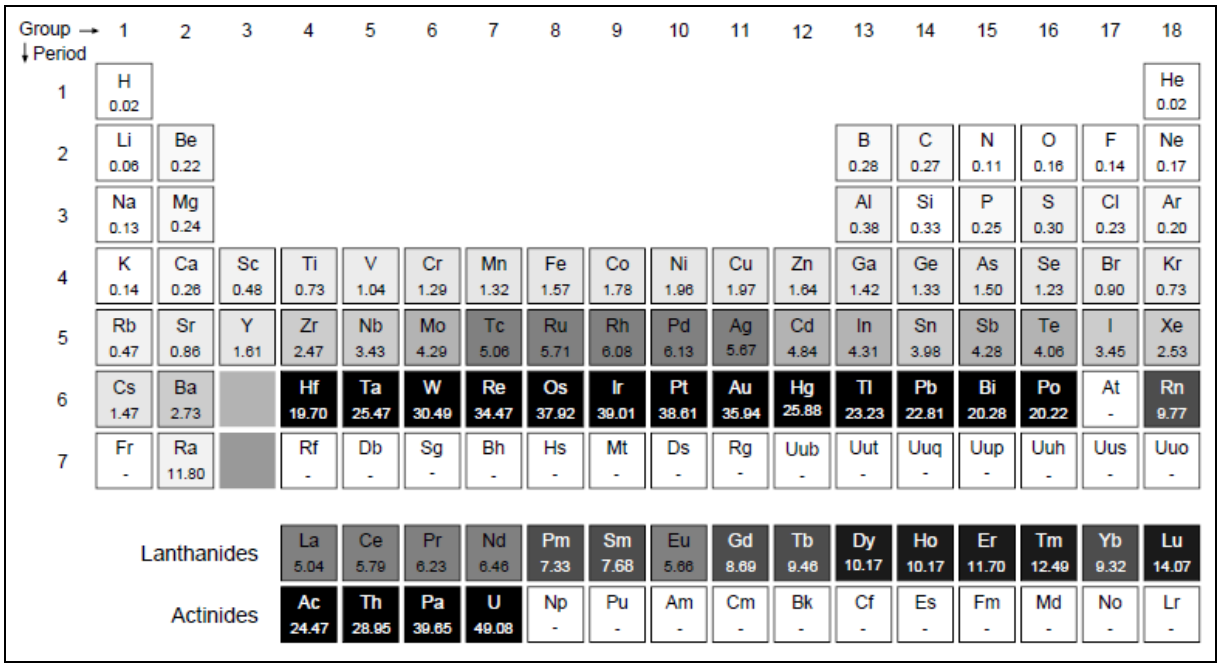


Fig. 3: *Mass-attenuation coefficients for Al of X-rays in the energy range between 1 keV and 100 MeV; source NIST [2]*

Neutrons (and in particular thermal and cold ones) interact only with the atomic nuclei and ignore the electrons in the shell completely. They can either be captured or scattered (in some few cases they induce fission). The interaction probability with matter is depending not only on the elements but also on the special isotope because the interaction with the nuclei can be very specific. For natural elemental compositions the attenuation probability the situation is shown in Fig. 4 – in comparison to the similar X-ray data.
It has to be highlighted that neutron interaction probabilities (described by the attenuation coefficients) does not underlie a systematic law when different materials are compared. There are prominent neutron absorbing isotopes (He-3, Li-6, B-10, Gd-157, …) whereas some other materials only weakly interact with neutrons (Pb, B-11, He-4, Zr). The strong absorbers are used as materials for neutron detection (see below).

In this respect, it depends on the material composition and the sample dimensions which radiation (X-rays or neutrons) have to be taken for the investigation best. The numbers in Fig. 4 can be taken directly to check the transparency of the material layer under investigation. According to (3), they describe to which intensity the beam drops down after passing a sample with the thickness of 1 cm.



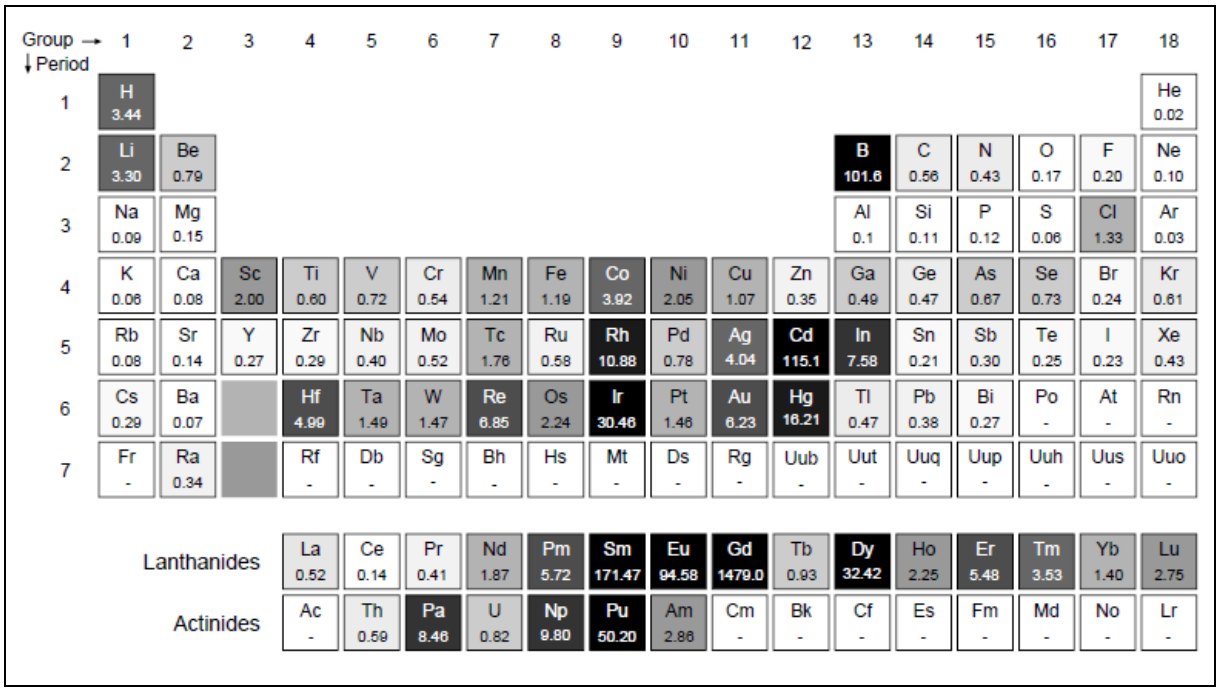


Fig. 4: *Attenuation coefficients for X-ray (100 keV) – above - and thermal neutrons – below – in the unit cm^{-1} . While a systematic increase of the attenuation contrast occurs for X-rays, no simple and determined relation to the mass number can be seen for neutrons. It is the strength of neutrons to provide a high contrast for hydrogen and other light elements and a high transmission ability for heavy elements.*

3 Application Range of transmission imaging techniques (example wood)

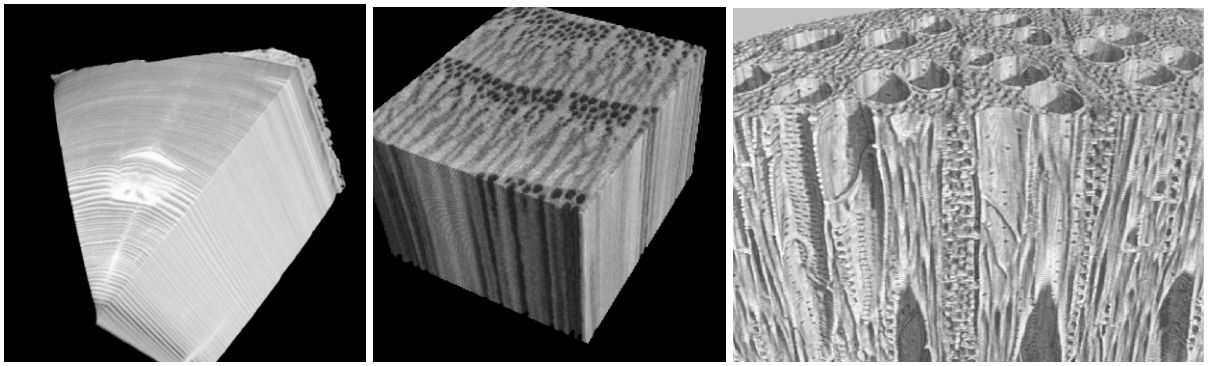


Fig. 5: *Tomographic views on wood samples on different size scale, obtained with X-rays (100 keV) – left; cold neutrons – middle; synchrotron light – right. The dimensions of the samples are – from left to right: about 10 cm, 1 cm, 1 mm.*

Although all three transmission techniques are performed at completely different facilities using at least two different kinds of radiation the findings about material properties can be quite specific in all three cases. This should be illustrated by a common material: wood.

As demonstrated in Fig. 5 with wooden samples of different sizes, structural information can be obtained with tomography methods (see below) in all shown 3 cases; however the details in respect to sample size and spatial resolution vary over 3 orders of magnitude. This behaviour is mainly caused by the attenuation of the applied beams given by the sample properties in respect to that radiation. In addition, the beam size and the inherent detector resolution contribute to the image features. While the attenuation coefficient for wood of X-rays of 100 keV is only 0.1 cm^{-1} it reaches 1.3 cm^{-1} for thermal neutrons. Accordingly, smaller samples with higher resolution and adequate contrast can be observed with neutrons. It has to be mentioned, that the interaction of neutrons with wood is mainly given by hydrogen, whereas in the case of X-ray mainly carbon and oxygen contribute to the image contrast and hydrogen plays nearly no role.
Therefore, neutrons are preferentially be used to determine the water content in wood with high sensitivity and accuracy [3].
In the case of synchrotron light, structural information can be obtained with a resolution similar to that for electron microscopy – however in 3D and without specific sample

preparation. Here the contrast is obtained by choosing photon energies of only few keV where the attenuation coefficients reach higher values too (see e.g. Fig. 3)

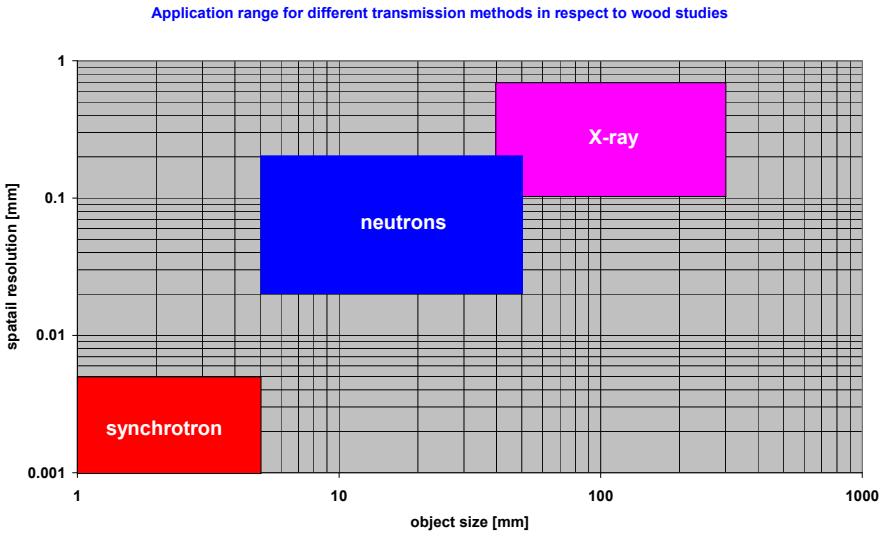


Fig. 6: *The use of X-rays, neutrons and synchrotron radiation can cover a wide range of object size. This behaviour is given by the attenuation properties of wood in respect to the applied radiation, the beam conditions and the detector properties.*

4 Imaging Facilities

As shown simplified in Fig. 2, the setup for imaging consists of the source and the detector (behind the object). In reality, the equipment has to be surrounded by a suitable shielding as all three kinds of radiation have certain impact to human health depending on the dose. While X-ray imaging facilities can be found with different size and source strength in laboratories, neutrons and synchrotron radiation for imaging purposes have to be produced by large scale facilities. PSI (Switzerland) is operating two such units: SINQ – a steady state spallation neutron source and SLS – a third generation synchrotron light source. Both are equipped with imaging beam lines as it will be shown below.

4.1 X-rays

The generation of X-rays has been done since Röntgen's discovery according to the same principle: electrons are accelerated within an electric field of high voltage (some kV) and stopped on a metallic plate, the anode. Photons are emitted about perpendicular to the electron beam direction with energies below that high voltage level such counted in the unit keV (see the principle in Fig. 7 on the left). Beside this so called "Bremsstrahlung", photons with a characteristic energy are emitted too which corresponds to the levels in the different atomic shells of the anode material when an electron is removed from that shell and replaced by another one, accompanied by the photon emission. A typical photon emission spectrum is shown in Fig. 8.

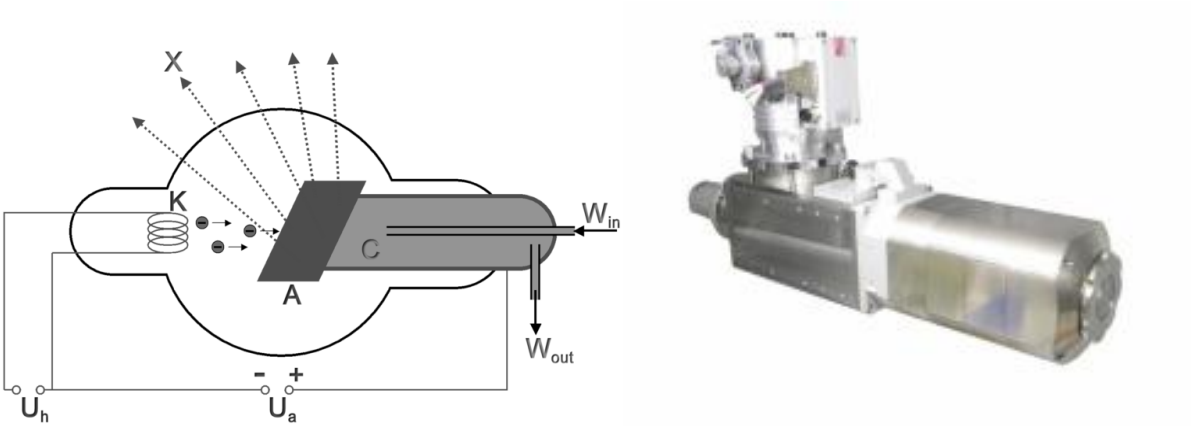


Fig. 7: *The principle of an X-ray tube (left): electrons are accelerated in an electrical field and photons with energies in the keV range are emitted from the anode. A state-of-the-art micro-focus tube is show on the right.*

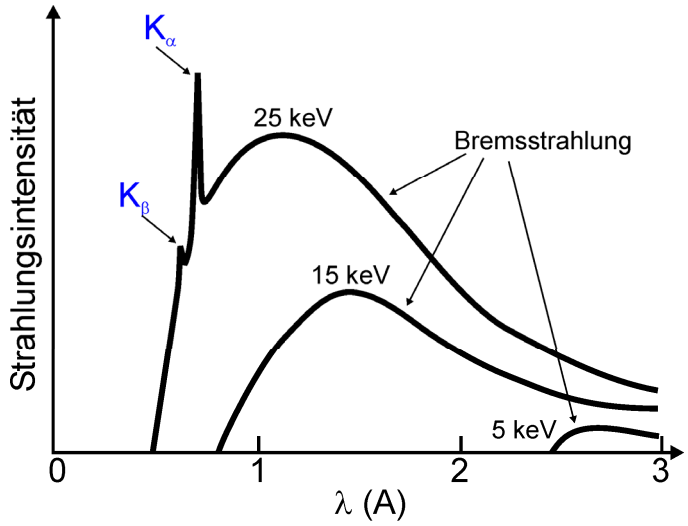


Fig. 8: *The emission spectrum of an X-ray tube consists mainly of the "Bremstrahlung" – a continuous spectrum until limits given by the high voltage. The characteristic lines exited in the anode material are overlaying the continuous part.*

X-ray imaging has been very common in medical diagnostics in the radiography mode since Röntgen's days until today. X-ray films have been used routinely where soft tissue occurs transparent and heavier parts of the body (e.g. bones) deliver higher contrast.
The step to digital X-ray imaging by using new types of image detectors enables nowadays also methods like tomography, laminography or real-time imaging.
X-ray imaging is used today more and more as research tool and as method for non-destructive testing in industry. Here, different and higher demands for advanced imaging systems came up. In particular for metallic objects, higher photon energies are required and tube voltages up to 420 kV become quite common.

On the other hand, images with much higher spatial resolution are possible if options for image magnification are taken into account. Using a cone beam setup (see Fig. 2, right), the small point source enables manifold magnification when the object is set close to the source, but the detector further away from the object. Finally, the resolution is given by the focal spot of the source which can be in the order of few micro-meters only. These tube systems are highly sophisticated as the source power and the heat removal have to be balanced in the right way.
Even if the advanced micro-systems have reached a very high level in spatial resolution they can not fully compete to the imaging systems at synchrotron light sources. In particular, their source strength is orders of magnitude smaller compared to the synchrotron sources ending up in much longer acquisition times.

4.2 Neutrons

Neutron imaging is much less available than that with X-rays since a beam line at strong neutron source is required. Either fission reactors or spallation neutron sources have been used for this purpose while radioactive sources are much too weak for practical use. A few accelerator devices started with neutron imaging capabilities. Due to the small number of such strong sources and the limited access to beam ports, there are only 10 to 15 competitive neutron imaging beam lines available world-wide (see Appendix).
In respect to image contrast, transmission and image quality, thermal or cold neutrons have been found most useful. Fast neutrons are less common since they provide inherent blurring in the detector even if they can transmit often thicker layers of materials.
A typical neutron imaging beam line layout is given schematically in Fig. 9. The fast neutrons from the initial source have to be slowed down in a moderation process by either hydrogen or deuterium before a part of the thermalized neutrons is extracted and guided to the sample. The transmitted neutrons are either measured in the detector behind the object or stopped in the beam dump.

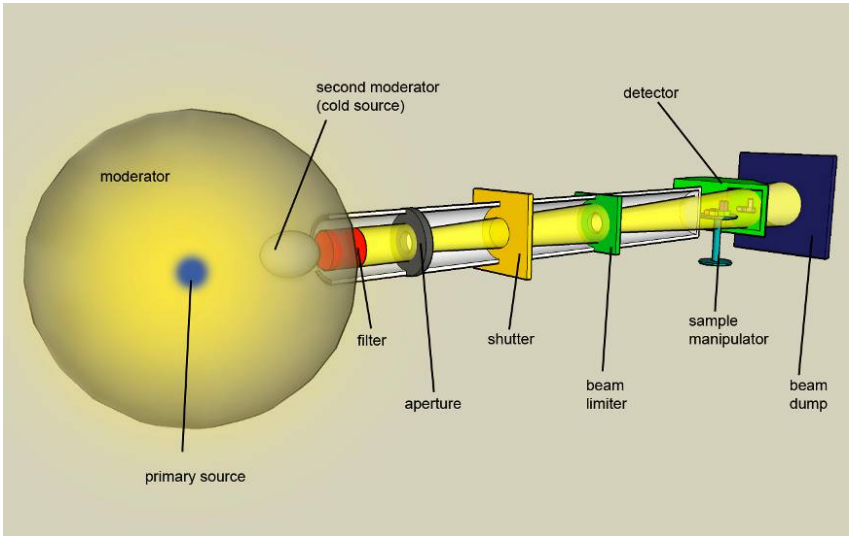


Fig. 9: *Layout of a beam line for neutron imaging describing the different components: the length is typically in the order of 10 meters, the beam size about 20 cm and can be varied.*

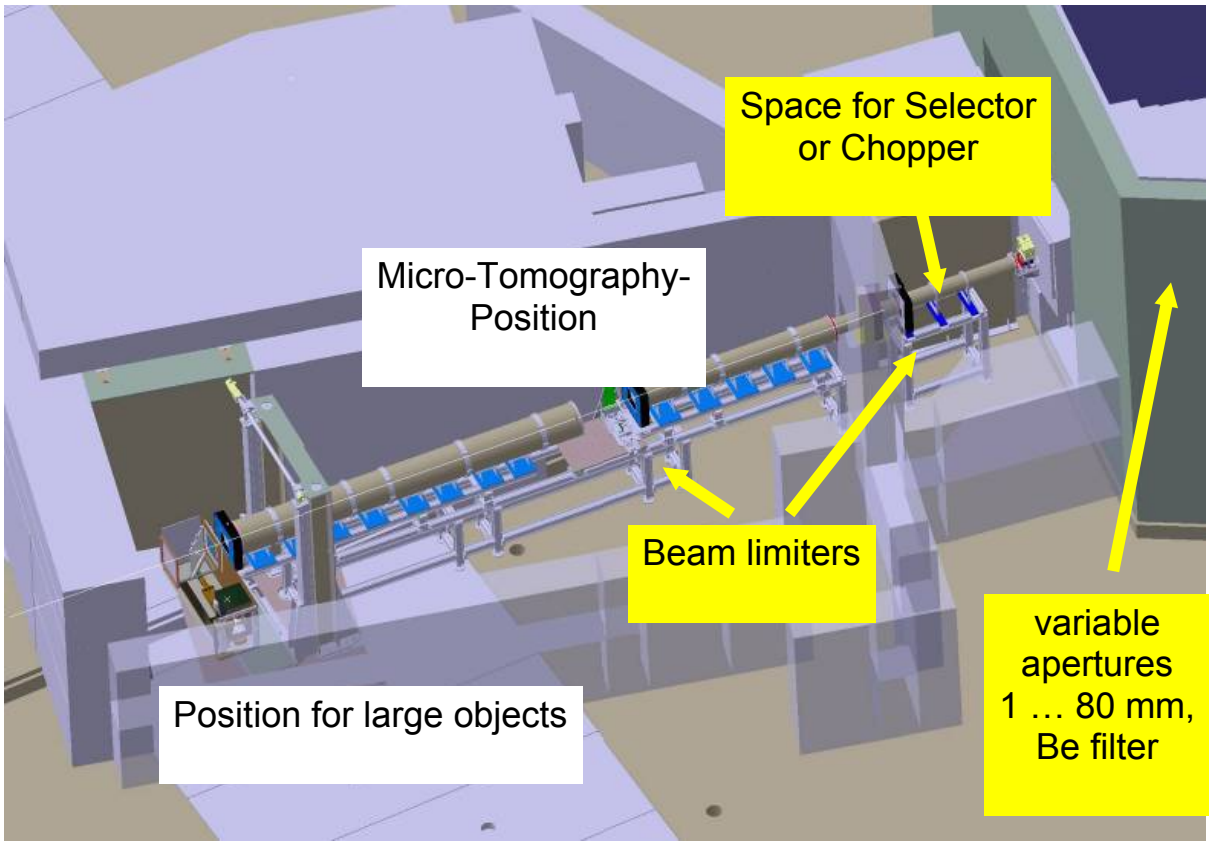


Fig. 10: *The ICON facility at SINQ (PSI, Switzerland) for imaging with cold neutrons; the length of the beam tube is about 12 meters; the aperture can be changed remotely controlled in steps from 1 mm to 80 mm.*

Based on this principle a facility as shown in Fig. 10 with the example of ICON at PSI can be realized. It should be built quite flexible in order to satisfy the demands of the different facilities users in respect to beam size, intensity and divergence. Accordingly, different detection systems have to be installed and used for the specific applications (see below).

4.3 Synchrotron radiation

Synchrotron light is emitted when charged particles (preferential electrons) with high energy are caused to change their direction by means of magnetic fields. Photons with keV energies are sent out into the initial direction of the charged particles. This previously misleading effect has been used to build dedicated large ring accelerators for electrons hosting many beam lines for X-rays aligned in tangential direction to the annular electron accelerator.
The intensity of such extracted beams of X-rays is by orders of magnitude higher in comparison to X-ray tubes. It can be tuned in several respects: collimation, coherence, spot size, energy band and intensity.
Facilities for imaging purposes are placed at several synchrotron light sources (see Appendix). As an example, the layout of the TOMCAT beam line [4] at SLS, PSI, CH is

shown in Fig. 10. The high intensity of the beam can be used either to create images with superior spatial resolution (far below 1 μm) or obtain image sequences (also in tomography mode) with extremely high frame rate. The beam size is in the order of few mm only which limits to applications of microscopically small samples (e.g. Fig. 5 – right).

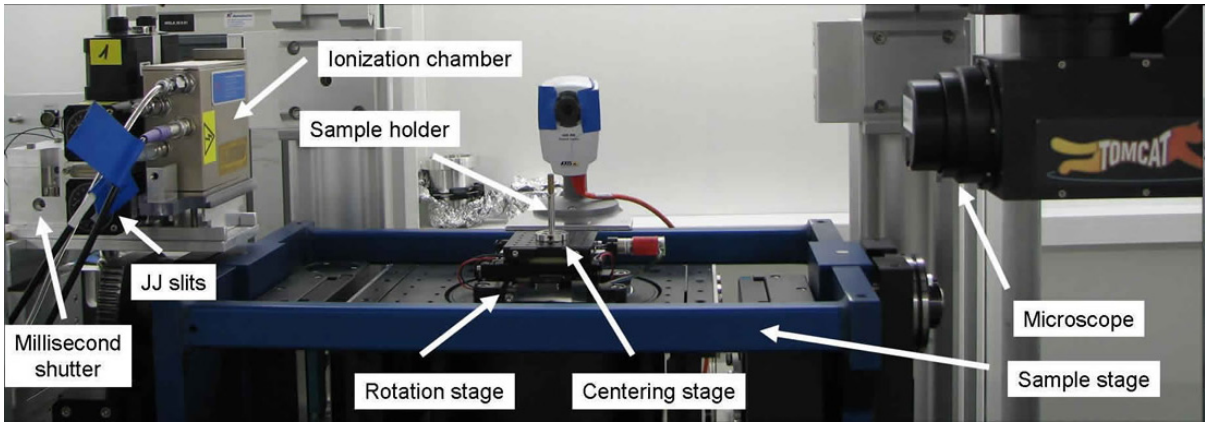


Fig. 11: *The TOMCAT imaging facility at the SLS synchrotron light source [4] (PSI, Switzerland) with some of the components described*

4.4 Imaging detectors

Film methods were used exclusively in the very beginning of transmission radiography. Although its inherent resolution is quite high there are many limitations in respect to the dynamic range, the linearity, the sensitivity and the missing options for data treatment. In the 90ies of last century, several digital imaging systems were invented and implemented. As primary radiation detection system scintillator screens, semi-conductor arrays or imaging plates have been used [5].

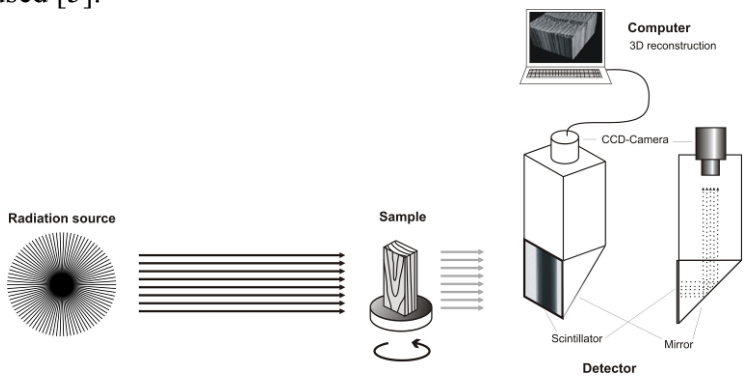


Fig. 12: *The setup of a digital imaging system based on a CCD camera observing the light from a radiation sensitive scintillation screen; under modified and specially arranged conditions this setup is in use for neutrons and X-rays as well*

Relatively simple, but flexible and most efficient is the combination of a highly sensitive camera (mainly CCDs, now also CMOS) with a radiation sensitive scintillation screen. A

typical setup is shown in Fig. 10, where the light from the scintillator is reflected out of the direct beam towards the camera, which is protected from direct irradiation in this manner.
Depending on the beam intensity, the exposure time with camera based systems is in the order of only seconds compared to minutes when a film has to be exposed. The result of imaging with digital systems is a data set containing a pixel matrix of grey values. State-of-the-art systems work with 16 Bit dynamic range enabling to distinguish more than 65.000 grey levels.
X-rays are able to excite detection systems mentioned above directly. This is not the case for neutrons due to their missing electric charge. A primary ionisation in the detector material is not possible therefore. Neutron detection has to be done via a nuclear reaction, mainly neutron capture. For this purpose, the following materials are in use mainly: Li-6, B-10, Gd and He-3. The reaction products – ions, α-particles, conversion electrons or gamma radiation – induce in a second step the measurable signal. For neutron imaging most common are scintillators in a combination of Li-6F mixed with ZnS [6].
Semiconductor arrays as pixel detectors [7] or amorphous silicon flat panels and imaging plates are in use as imaging devise beside camera detection systems. More advanced systems [8] are able to cover a high spatial resolution and high temporal resolution which can be used for an energy selection in the time-of-flight mode (see below).

5 Imaging techniques

The imaging with all three kinds of radiation has similarities and specific differences. In the "normal" radiography mode, the obtained image is an overlay of all material information in beam direction. For flat slabs there is no problem to distinguish the material distribution which is set "transparent" by the transmitting radiation. An example is given for the case of a hard disk drive with a more flat structure. Characteristic differences between X-ray and neutron imaging can be found whereas the inherent image quality is about the same.

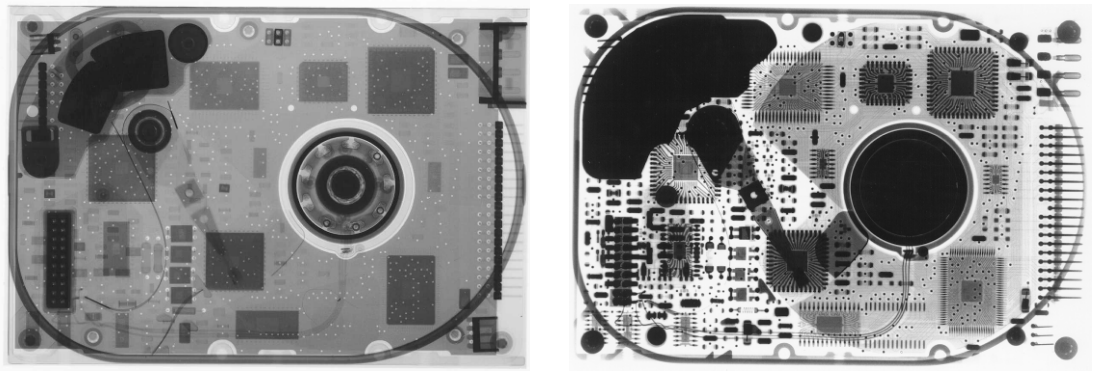


Fig. 13: *Neutron transmission radiography image (left) compared to an X-ray image (right) of a hard disk drive: metallic parts are more transparent for neutrons whereas a higher contrast is obtained for the organic material of the conductor board.*

5.1 Tomography

The third dimension of an object under investigation can be explored with the help of tomographic methods. For this purpose, information has to be derived from different orientation of the object, that means projection have to be taken from different viewing angles while the object is rotated around a well aligned vertical (or horizontal) axis.
Meanwhile the reconstruction methods, mainly based on filtered back-projection and implemented in several software tools are able to convert the projection data via sinograms into the volume data of the object as a matrix of voxel values in the format of attenuation coefficients (see (4)). All reconstruction algorithms are based on the assumption of the validity of that law over all dimensions of the object what is true only in first order.
If the volume data are available, software tools for visualisation and evaluation can be used for arbitrary virtual slices through the object and the separation of zones with the same attenuation contrast in segmentation routine. An example is given in Fig. 14 for a Buddha sculpture from the 14th century imaged in neutron radiography and tomography mode. In this case, the inner organic content (wooden stick, a wire, textiles, plants and seeds) can be visualized within the metallic cover – a case where neutrons exclusively have to be applied [9].

Fig. 14: *Comparison between a transmission neutron radiography image (middle) of a Buddha sculpture (Buddha Sakyamuni) from the 14th century (left) with a virtual slice derived from the tomography data set (right).*

Nowadays, tomography techniques are common with all three considered techniques. Based on the considerations in chapter 2 and 3 it has to be decided which method has to be used in respect to the required spatial resolution and its contrast. Compared to a simple transmission image, tomography requires about 300 – 500 times more acquisition time and a lot more effort for volume reconstruction and data visualization. The data to be stored and managed are in the GBytes region and a suitable data handling system is required.

5.2 Time-dependent studies

The study of processes in the time domain is very important with transmission radiation which can visualize inner structures and their sequential development. Generally, it depends on the beam intensity which degree of time resolution can be obtained. In this respect, synchrotron radiation is a favourite to provide the best conditions.
In respect to neutron imaging, the limitation to flux levels in the beam of 10^8 to 10^9 $cm^{-2}s^{-1}$ in best cases is a real handicap. Considering a pixel size of 0.5 mm * 0.5 mm as reasonable and assuming a 100% detector efficiency for the flux of 10^9 $cm^{-2}s^{-1}$ the time per frame cannot be better than 0.4 milliseconds when about 1000 counts are needed for a valid contrast. This performance can be obtained on only few places word-wide. In most cases, acquisition times of several seconds per frame are common. Beside the acquisition time, the read-out from the detector is quite critical and can take several seconds too.
For real applications there are two types of approaches:

1. Fast sequences of frames with limited image quality which is given by the compromise between the acquisition time and the detector noise. The typical frame rate (for neutrons) is in the order of 3 to 50 fps, depending on the beam intensity.
2. Stroboscopic imaging of repetitive processes where a stack of partial images is build to reach the best image quality from a very narrow time frame (e.g. 10 µs) – integrated over several thousands of cycles. An example is given in Fig. 15 for a running combustion engine under full load on 8000 rpm.

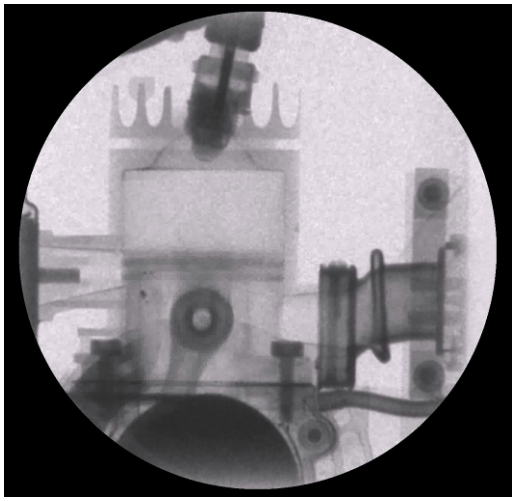


Fig. 15: *Investigation of the operation within a combustion engine in the real-time mode where the repetitive process is imaged through the metallic structure by means of neutrons. The engine was fired and operated on 8000 rpm, whereas a single frame was taken within 50 µs. Stacking about 1000 such identical images delivered the valid and useful information.*

Future application will be found at pulsed neutron sources where the neutron spectrum is spread over a certain range in the time domain, which brings us to energy-selective studies (see below).

5.3 Energy selection

The radiation sources emit primarily a brad spectrum of neutrons and X-rays, respectively. It is an advantage to apply only a narrow energy band of that radiation because the specific material properties can be enhanced. The attenuation cross-sections have quite strong energy dependence as shown in Fig. 2 for X-rays and for some structural materials in the energy range of "cold neutrons" (Fig. 14). The application of a broad spectrum of radiation averages and smooths the specific desired material behaviour in some respect and cases.

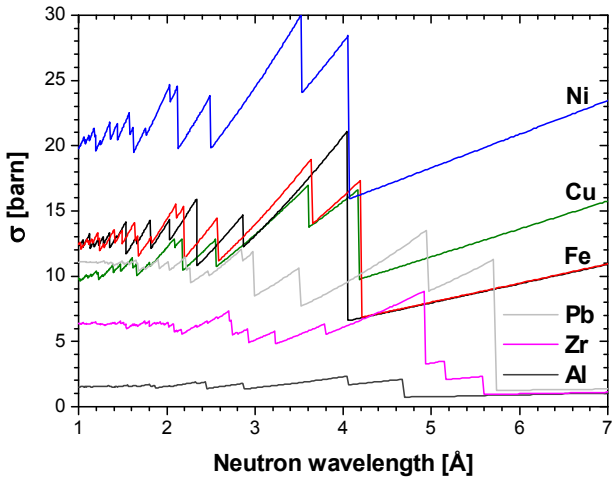


Fig. 16: *Microscopic cross-sections of crystalline structural materials in the cold neutron energy range are showing the characteristic Bragg edges caused by the neutron scattering at crystal lattice planes.*

Energy selective investigation can enhance the contrast between different materials when the right energy is chosen close to one or the other Bragg edges. Image referencing between two images at the same spatial position but taken with two different energies can further enhance the contrast by image post-processing techniques.
Because the Bragg behaviour is caused by the crystalline structure of the investigated materials the orientation and crystal distribution can be visualized directly in cases when the crystal size fits to the spatial resolution of the imaging system. Such an example is given in Fig. 17, where the laminar structure of a rolled Al sample becomes visible. At least three zones with different crystal structure (not composition!) can be distinguished. In poly-chromatic mode and in the absorption range above 5 Å the sample looks quasi-homogenous. This method can be used for a pre-characterisation of unknown materials in a simple and non-destructive way.

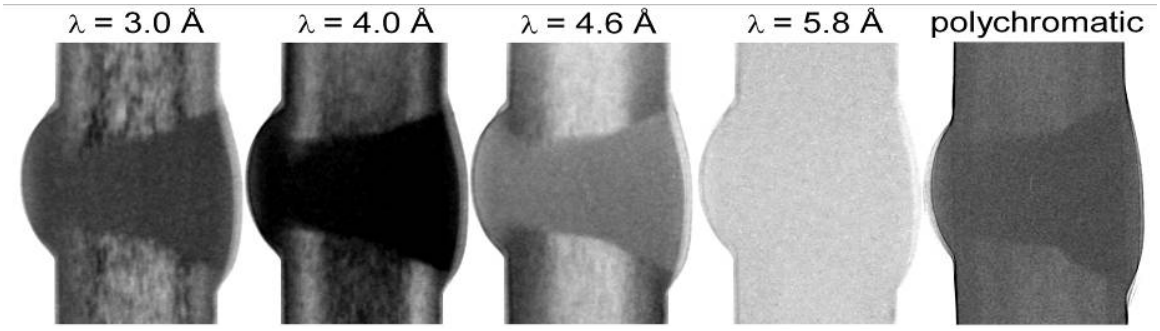


Fig. 17: *Part of a rolled Al plate with a weld in the middle was studied at different neutron energies and structural differences can be derived in the plate but not in the weld where the grain size is too small compared to the detection limits.*

Presently used devices for energy selection (turbines, mono-chromatizers) are able to cut out energy bands of 5% to 15%, depending on the specific performance parameters. With the help of the upcoming pulsed spallation neutron sources the time-of-flight options have much less limitations and will allow the selection of narrower band width.

5.4 Phase contrast imaging

Nearly all results of investigations described above were based on attenuation in the absorption contrast mode even if scattering of neutrons or photons might play a major role. In this way, all radiation which is lost in the direct beam behind the object is handled in the attenuation law equally.

However, neutrons have also wave properties like X-rays. In the case of the neutrons, their wavelength λ can be calculated according to de Broglie's formula:

$$\lambda = \frac{h}{m \cdot v} \tag{4}$$

with h as the Planck's constant, m the neutron mass and v its velocity. One important feature of the interaction of waves with matter is refraction. The wavelength is changed as in classical optics and an index of refraction n (ratio between the velocities inside and outside the object) can be derived to depend on the wavelength and the cross-section of the coherent scattering:

$$n = \frac{v_i}{v_0} = 1 - \delta = 1 - N \cdot \sigma_{coher} \cdot \lambda^2 / 2 \cdot \pi \tag{5}$$

In reality, δ is in the order of 10^{-5} only and can be either positive (e.g. for Al, Cu, Fe) or even negative (e.g. for Ti or Mn).

This little phase shift can be used in practical imaging to enhance the contrast at edges in order to distinguish materials with the same attenuation behaviour but different refraction properties. Detection systems with high spatial resolution are needed to take profit from this behaviour. An example from neutron edge enhancement by refraction is shown in Fig. 18.

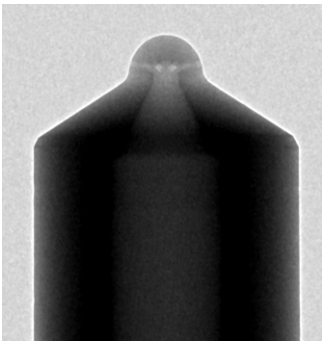


Fig. 18: *Upper part of a Diesel injection nozzle observed with cold neutrons and a high resolution imaging detector; the edge enhance effect is clearly visible caused by the refraction in steel; sample diameter 4 mm.*

Beside that edge enhancement, it was possible to derive the differential phase information by special experimental setups with a grating interferometry technique [10] as an additional material parameter. This technique is now available in X-ray systems and at cold neutron beams as well.

6 Applications

Practical applications in transmission radiation imaging for non-invasive investigations of objects are progressively in use in different fields of science and engineering. It is the increasing availability of X-ray systems in many labs and the strong improvement of digital imaging systems, including the computation and evaluation techniques which push this attractive field forward. Based on considerations about the attenuation behaviour of the sample material, the demands in resolution and the dimensional boundary conditions the best suitable setup has to be chosen.

6.1 Material research: metals, bio, geo-science

The complementarities between X-rays and neutrons were roughly explained in chapter 2. The consequence for practical applications can be seen e.g. in Fig.13: neutrons have in most cases higher penetration ability through metals while a strong contrast is given for hydrogenous materials even in a metallic cover. Therefore thin parts of an animal body can well be visualized in neutron tomography (see Fig. 19).
Unfortunately, the resolution in neutron imaging is presently limited to about 20 μm although the contrast might be high enough to see more details in biological samples. Studies with demand in higher resolution should be done with synchrotron radiation even the contrast might be quite low.
Studies with geological samples gained high interest because structural information on different size scale can be derived with tomography methods. An important feature is the study of the porosity and the interaction with water flow phenomena. Here, neutrons can provide high contrast even in small moisture contents. Finally, the behaviour of plant roots in soil were studied with success during the feeding with moisture and nutrients [11].

Fig. 19: *Tomographic views onto a fly which was studied using cold neutrons with presently highest possible resolution (pixel size 13.5 µm; sample size12 mm); the advantage of the high contrast for hydrogen enables to see very much detail – better than naked human eye can do.*

6.2 Industrial applications

Several industrial companies have started to install X-ray imaging systems for the inspection and quality insurance next to production lines or in dedicated in-house labs. Only if these installations fail to reach the inspection goals other techniques like neutron imaging or imaging with synchrotron radiation is taken into account. In these cases, the industrial partner needs to enter a large scale facility and has to book beam time on commercial basis.

In respect to neutron imaging there is a short list of prominent applications without any guarantee for completeness:

- Polymer Electrolyte Fuel Cells
- Combustion engine operation, fuel injection
- Batteries (Lithium-ion)
- Diesel particular filters
- Adhesive and brazing connections
- Welds in structural materials
- Explosives
- Two-phase flow processes
- Fuel and cladding for Nuclear Power Plants

There will be found several new applications for the different transmission imaging techniques, based on the permanently changing and increasing demands by industry. A high degree of flexibility is required in order to satisfy the demand in the right manner by further improvements in the experimental techniques.

6.3 Cultural heritage studies

The demands from museums and archaeologists to study and to visualize hidden structures and features in objects of cultural heritage value are mainly coupled with the need for a non-destructive or non-invasive approach because no damage is allowed. There is a tradition to

investigate museums objects with X-rays [12], but this cannot solve all request due to missing transmission, limited contrast and spatial resolution. Therefore, the availability of new imaging techniques with neutrons, synchrotron radiation and by using new imaging methods is very welcome by the art historical scene.
In the framework of this course only a few examples (e.g. Fig. 14) can be given. In some cases, the X-ray and neutron data fits well together to create a comprehensive picture of the object. In some other cases, only the one or other method can solve the task to visualize the important content. In the case in Fig. 14, X-ray would completely fail to see the inner content of the sculpture due to missing contrast for the organic content and limited transmission through the metallic cover of the statue.

7 Summary and Outlook

Transmission imaging techniques based on X-rays or neutrons are modern tools for research and technical applications. An enormous gain in performance happened in the last few years by the introduction of advanced digital systems, the implementation of sophisticated methods like phase imaging and tomography on routine operation base.
Already with table-top laboratory installations for X-ray tomography the resolution came close to the micro-meter region.
The highest possible performance can still be found at large scale facilities like synchrotron light sources and strong neutron sources. Their beams can be used to create highest resolution images, fast image sequences, energy dispersive images or images with superior dynamic and low noise.
However, there is still potential for further improvements by the implementation of better detection systems, methodical developments and the extension of the user application range. In this respect, the combination of imaging with diffraction and the direct combination of neutron with X-ray imaging are promising approaches. Another future trend will be the improvement in the energy resolution in neutron imaging my means of time-of-flight methods at pulsed sources.

Appendices

Facilities for transmission imaging with neutrons and synchrotron radiation world-wide

Country	Location	Institution	Facility	Neutron Source	thermal/cold flux [cm^{-2} s^{-1}]	L/D - ratio	Field of View
Austria	Vienna	Atominstitut	imaging beam line	TRIGA Mark-II, 250 kW	1.00E+05	125	90 mm diam.
Brazil	Sao Paulo	IPEN	imaging beam line	IEA-R1M 5 MW	1.00E+06	110	25 cm diam.
Germany	Garching	TU Munich	ANTARES	FRM-II 25 MW	9.40E+07	400	32 cm diam.
Germany	Garching	TU Munich	NECTAR	FRM-II 25 MW	3.00E+07	150	20 cm diam.
Germany	Berlin	HZB	CONRAD	BER-II 10 MW	6.00E+06	500	10 cm * 10 cm
Hungary	Budapest	KFKI	imaging beam line	WRS-M 10 MW	6.00E+05	100	25 cm diam.
Japan	Osaka	Kyoto University	imaging beam line	MTR 5 MW	1.20E+06	100	16 cm diam.
Japan	Tokai	JAEA	imaging beam line	JRRM-3M 20 MW MTR	2.60E+08	125	25 cm * 30 cm
Korea	Daejon	KAERI	imaging beam line	HANARO 30 MW	1.00E+07	190	25 cm * 30 cm
Switzerland	Villigen	PSI	NEUTRA	SINQ spallation source	5.00E+06	550	40 cm diam.
Switzerland	Villigen	PSI	ICON	SINQ spallation source	1.00E+07	350	15 cm diam.
USA	PennState Uni.	University	imaging beam line	TRIGA 2 MW	2.00E+06	100	23 cm diam.
USA	Gaithersburg	NIST	CNR	NBSR 20 MW	2.00E+07	500	25 cm diam.
USA	Sacramento	McCleallan RC	imaging beam line	TRIGA 2 MW	2.00E+07	100	23 cm diam.
South Africa	Pelindaba	NECSA	SANRAD	SAFARI-1 20 MW	1.60E+06	150	36 cm dia.

A: Prominent facilities for neutron imaging around the world

Country	Location	Institution	Facility	Source	Energy
Switzerland	Villigen	Paul Scherer Institut	TOMCAT	SLS	12-40 keV
USA	Upton	BNL	XCMT	ALS	6-36 keV
Germany	Karlsruhe	FZ Karlsruhe	TopoTomo	ANKA	> 6 keV
France	Grenoble	ESRF	ID22	ESRF	
Germany	Hamburg	DESY	P06	HASYLAB	
UK	Didcot	Rutherford Lab	I13L	DIAMOND	6-30 keV
Japan	Harima Science Park City	RIKEN	BL 20	SPRING 8	
USA	Chicago	ANL	5-BM-C and others	APS	
France	Saclay	France joint effort	PSICHE	SOLEIL	15-100 keV
Italy	Trieste	AREA Science Parc	SYRMEP	ELETTRA	

B: Prominent facilities for imaging with synchrotron radiation around the world (selection)

References

[1] RR data base IAEA: http://nucleus.iaea.org/RRDB/RR/ReactorSearch.aspx
[2] www.nist.gov/pml/data/xraycoef/index.cfm
[3] D. Mannes, Non-destructive testing of wood by means of neutron imaging in comparison with similar methods, Diss. ETH Zurich No. 18563, 2009
[4] www.psi.ch/sls/tomcat/beamline-layout
[5] E.H. Lehmann, A. Tremsin, C. Grünzweig, I. Johnson, P. Boillat and L. Josic, Neutron imaging — Detector options in progress, 2011 JINST 6 C01050
[6] G. Frei, E.H. Lehmann, D. Mannes, P. Boillat, The neutron micro-tomography setup at PSI and its use for research purposes and engineering applications, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 605, Issues 1-2, ITMNR 08 - Proceedings of the sixth Topical Meeting on Neutron Radiography, 21 June 2009, Pages 111-114, ISSN 0168-9002
[7] E. H. Lehmann, P. Vontobel, G. Frei, C. Br.onnimann, Neutron imaging—detector options and practical results, Nuclear Instruments and Methods in Physics Research A 531 (2004) 228–237
[8] A.S. Tremsin, J.B. McPhate, E.H. Lehmann, J.V. Vallerga, O.H.W. Siegmund and W.B. Feller, High resolution neutron radiography with very compact and efficient neutron collimators, 2011 JINST 6 C01041
[9] E.H. Lehmann, S. Hartmann, M.O. Speidel, Investigation of the content of ancient Tibetan metallic Buddha statues by means of neutron imaging methods, Archaeometry 52, 3 (2010) 416-428
[10] Pfeiffer F., Gruenzweig C., Bunk O., Frei G., Lehmann E., and David C. Neutron Phase Imaging and Tomography PHYSICLAL REVIEW LETTERS 96, 215505, June 2, (2006)
[11] A. B. Moradi, A. Carminati, D. Vetterlein, P. Vontobel, E. Lehmann, U. Weller, J. Hopmans, H.-J. Vogel, S. Oswald, Three-dimensional visualization and quantification of water content in the rhizosphere, New Phytologist (2011) 192: 653-663
[12] J. Lang, A. Middleton, Radiography of Cultural Material, Elsevier, 2005, ISBN 0 7506 6347 2

F 3 Photoelectron Emission Spectroscopy [1]

Claus M. Schneider
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 43rd IFF Spring School "Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources" (Forschungszentrum Jülich, 2012).

A Preliminary Note About This Lecture

Photoelectron spectroscopy has matured into an extremely versatile and powerful analysis technique. It permits access to a very wide variety of materials and their electronic structure, ranging from complex bulk structures down to free atoms. Consequently, there is an enormous wealth of results and interesting examples on different systems available, which are well worth being discussed. For good reasons, however, this lecture must focus on a few essential basics, novel aspects and a very personal selection of examples. For an in-depth study of photoemission spectroscopy and phenomena, the reader is referred to a number of excellent textbooks and review articles covering this field [1–5].

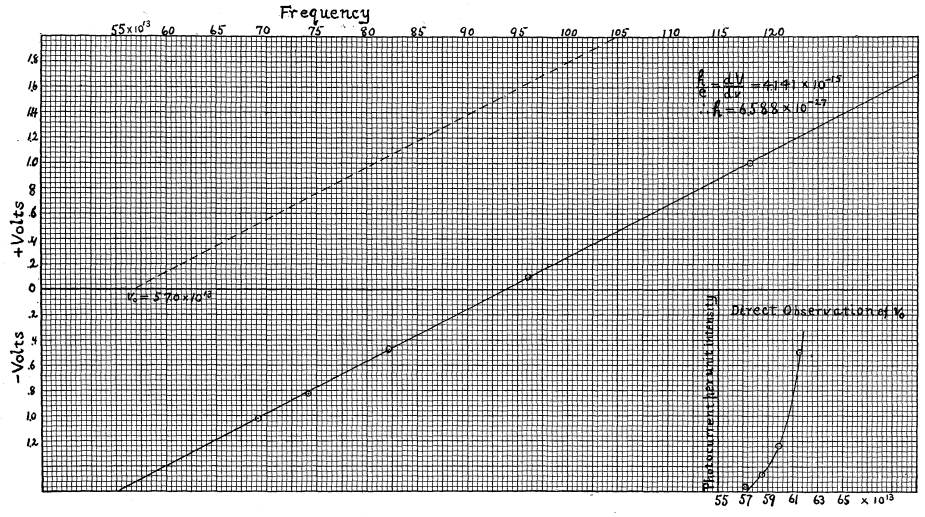


Figure 1: *Millikan's experiment on the photoelectric effect. For each exciting light frequency (data point) there is a well-defined maximum retarding voltage for the photoelectrons, the slope of the regression line yielding Planck's constant h. From [6].*

1 Introduction

The method of photoelectron emission spectroscopy goes back to the photoelectric effect, which was discovered by H. Hertz in 1887 [7] and refers to the phenomenon of electrons being ejected from a metal when illuminated by electromagnetic radiation. The explanation of the photoelectric effect by A. Einstein in 1905 [8] was one of the first triumphs of quantum mechanics. In his famous paper, Einstein extended Planck's quantum hypothesis by postulating that quantization was not a property of the emission mechanism, but rather an intrinsic property of the electromagnetic field. Using this hypothesis, Einstein was able to explain why the maximum kinetic energy E_{kin} of the emitted electrons varies with the frequency ν of the incident radiation as

$$h\nu = \Phi_0 + E_{kin} = \Phi_0 + \frac{1}{2}m_e v^2 \equiv eU \tag{1}$$

where h is Planck's constant, Φ_0 is a characteristic energy and called the work function, m_e, e and v denote electron mass, charge and velocity, respectively. This is exactly the result expected if photons are quantized with energies $h\nu$, and earned Einstein the 1922 Nobel prize in physics.

In the early photoemission experiments one basically determined the total photoelectron current $I_p(U)$ on a counter electrode as a function of a retarding voltage U. In this way, the maximum kinetic energy of the photoelectrons was determined from the condition $I_p(U) \Rightarrow 0$ (Fig. 1). From today's perspective this approach corresponds to an angle-integrated photoemission experiment [1].
The development of photoelectron spectroscopy started at the end of the 1950's with K. Siegbahn, who studied the energy levels of core electrons in atoms using excitation with x-rays [9]. Since the exact binding energy position of the core level depends on the chemical environment of the atom from which the photoelectron is emitted, Siegbahn coined the name Electron Spectroscopy for Chemical Analysis (ESCA) for this spectroscopic technique. He was awarded the physics Nobel prize in 1981 for his contributions to high-resolution electron spectroscopy.

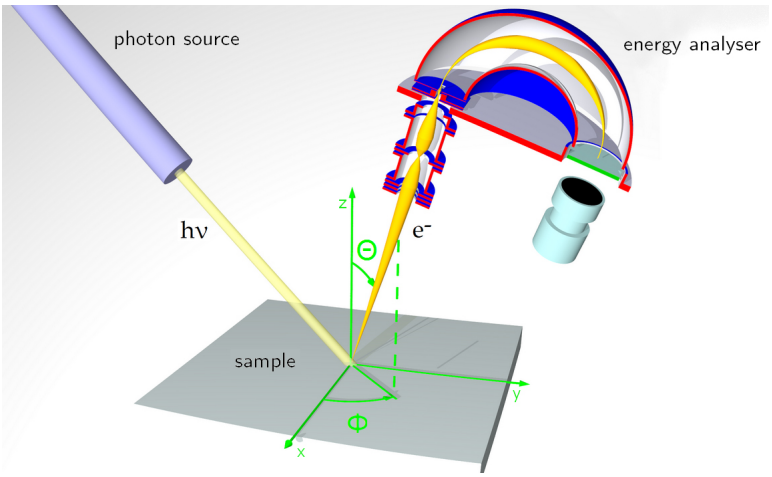


Figure 2: *Schematical picture of a modern photoemission experiment.*

A typical angle-resolved photoemission experiment as of today is sketched in Fig. 2. The photon source is typically a synchrotron radiation facility, which provides light over a broad range of photon energies $h\nu$ from the ultraviolet up to hard x-rays [10]. This light beam can be finely focused down to about 100 μm, its direction of incidence onto the sample and its degree and orientation of polarization (linear, circular) can be precisely controlled. The electron energy analyzer is equipped with an electron optical entrance lens system, which defines the angular range of the photoelectrons analyzed. Therefore, the emission direction denoted by the angles (θ, ϕ) is a free experimental parameter. On the one hand, for solid state experiments, mostly single-crystalline samples are investigated under ultrahigh vacuum conditions ($p < 10^{-8}$mbar). On the other hand, studies in catalysis require almost atmospheric pressure a the sample. This can be nowadays enabled by special designs of the entrance lens system involving differential pumping stages [11].

2 Fundamental Aspects of Photoemission

The valence electronic structure of a solid – in principle – can be calculated by solving a Schrödinger (nonrelativistic approximation) or Dirac equation (relativistic interactions included) for the respective lattice structure. For most metallic systems with weak electronic correlations, i.e. being close to the limit of a homogeneous electron gas, a quite successful description has been achieved within the framework of density functional theory (DFT) using the local density

(LDA) or more general local spin density approximation (LSDA) for the exchange-correlation potential [12]. In order to discuss the salient features of the photoemission process, we will first adopt this effective single-particle picture, although we must be aware of the fact that it does not capture electronic correlations properly.
One important result of the theoretical treatment is that the quantum mechanical wave function describing an electronic state in the solid depends on the symmetries of the Hamiltonian (e.g. lattice symmetries, inversion symmetry, time reversal symmetry, etc.) and must include also the electron spin. In the single particle picture the Hamiltonian can be written as

$$\left(\left[\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+eV(\vec{r})\right]+i\frac{e\hbar}{4m^2c^2}\vec{E}\cdot\vec{p}-\frac{e^2\hbar}{2mc}\vec{\sigma}\cdot\vec{B}- \right. \\ \left. -\frac{e\hbar}{4m^2c^2}\vec{\sigma}\cdot\left(\vec{E}\times\vec{p}\right)\pm eV_{exc}^{\uparrow\downarrow}(\vec{r})\right)\phi=E_{nls}(\vec{k})\phi. \tag{2}$$

with ϕ and $E_{nls}(\vec{k})$ denoting the single electron wave function and energy eigenvalue, respectively. The terms in square brackets in eq. 2 represent the Hamiltonian of a system subjected to an electromagnetic field (vector potential $\vec{A}$). This part contains all crystalline symmetries through the potential $V(\vec{r})$. The Darwin term ($\sim\vec{E}\cdot\vec{p}$) may be understood as a relativistic correction to the electron energy. The fourth term contains the interaction of the spins – described by the Pauli spin matrices $\vec{\sigma}$ – with an external magnetic field $\vec{B}$. The last two terms contain the spin-dependent interactions through spin-orbit coupling and the exchange-correlation potential $V_{exc}^{\uparrow\downarrow}(\vec{r})$. The latter is responsible for the formation of spontaneously ordered magnetic states in solids. All spin-dependent terms in the Hamiltonian tend to reduce the symmetry of the system in one way or the other, leading to the splitting and hybridization of degenerate states. A full set of energy eigenvalues obtained from the DFT treatment forms a band structure $E_n(\vec{k})$ of the solid with the band index n and the electron wave vector $\vec{k}$. The wave functions are Bloch functions and are further classified by the orbital momentum quantum number ℓ and the spin quantum number s. Formally, the respective states may be written as $|n,\ell,\vec{k},s\rangle$ with their energy eigenstates $E_{nls}(\vec{k})$.
In addition to this valence electronic states comprising delocalized electrons, the full electronic structure of a solid also contains atomic-like localized core levels at higher binding energies $E_B \gtrsim 30$ eV (the binding energy is referred to the Fermi energy, i.e. $E_B = 0$ at E_F). In the ground state, the core states are completely occupied, whereas the valence states are occupied up to the Fermi energy E_F in the case of metals. In semiconductors and insulators the Fermi level lies in a band gap and the intrinsic bulk states are occupied only up to the valence band edge in the undoped case. At semiconductors band bending and surface photovoltage phenomena occur, which complicate the intepretation of photoemission results from semiconductors [13]. They will not be treated in this contribution.
Irradiating such an electronic structure with photons leads to excitation of electrons from occupied into unoccupied states in the electronic structure. If these empty states are located above the vacuum level E_{vac} of the solid, photoelectrons can leave the crystal and can be measured by an electron spectrometer, yielding characteristic signatures of the valence electronic states and core levels (Fig. 3). The quantum mechanical essence of the photoexcitation process is captured in Fermi's Golden Rule, which describes the transition probability between two electronic levels $|i\rangle$ and $|f\rangle$ with binding energies E_i and E_f, respectively:

$$P_{i\to f}=\frac{2\pi}{\hbar}\left|\langle f|\,\mathcal{O}\,|i\rangle\right|^2\delta(E_f-E_i-h\nu) \tag{3}$$

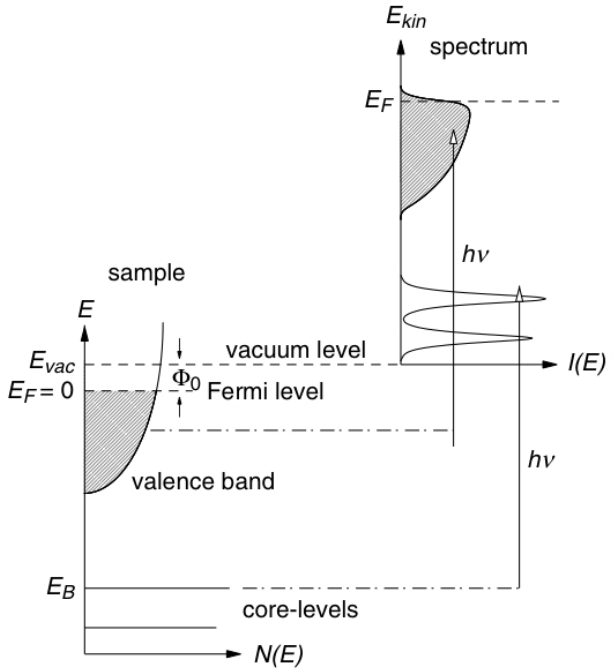


Figure 3: *Principle of the photoemission process. The electrons excited into states above the vacuum level form a photoelectron spectrum reflecting a broad valence electronic distribution (shaded area) and sharp emission lines from the core levels. From [3].*

In the simplest approach, the two levels $|i\rangle$ and $|f\rangle$ may be taken from the ground state electronic structure of the solid – which neglects the role of electronic correlations in the excitation process, as we will see below. The most important quantity in eq. (3) is the transition matrix element

$$M_{fi} = \langle f | \mathcal{O} | i \rangle \tag{4}$$

which essentially depends on the symmetries of the electronic wave functions and the photonic operator $\mathcal{O}$, whereas the delta function ensures energy conservation in the excitation process. For low photon flux densities the operator $\mathcal{O}$ can be treated within linear response theory and takes the form

$$M_{fi} = \frac{-e}{mc} \langle f | \vec{A}(\vec{r}) \cdot \vec{p} | i \rangle \tag{5}$$

with $\vec{A}(\vec{r})$ the vector potential of the electromagnetic field and $\vec{p}$ the momentum operator. It is usually assumed that the wavelength of the electromagnetic field is large compared to interatomic distances, i.e. $\vec{A}(\vec{r})$ varies only marginally in the spatial region contributing to the transition matrix element [1]. This view is commonly known as *dipole approximation* and simplifies the transition matrix element to

$$M_{fi} = \frac{-ie}{\hbar c} A_0 (E_f - E_i) \langle \psi_f | \vec{e} \cdot \vec{r} | \psi_i \rangle \tag{6}$$

with the complex amplitude of the vector potential A_0, its polarization vector $\vec{e}$, and the wavefunctions of the final and initial states ψ_f and ψ_i, respectively. This form of the transition matrix element is extremely valuable, as the quantity $\langle \psi_f | \vec{e} \cdot \vec{r} | \psi_i \rangle$ can be evaluated for selected symmetries of the wave functions and yields dipole selection rules, which are very useful for a

[1] This assumption should be revisited, if we go to photoexcitation with hard x-rays.

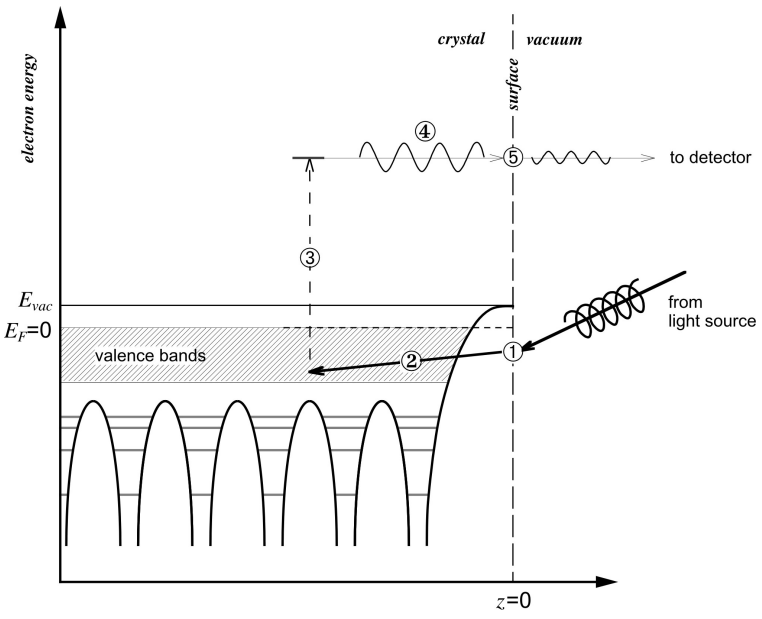


Figure 4: *Schematial representation of the three step model. The numbers denote: (1) refraction of the electromagnetic wave at the surface, (2) penetration of the photon into the solid, (3) photoexcitation, (4) propagation of the photoelectron to the surface, and (5) diffraction of the electron wave at the surface.*

qualitative interpretation of photoemission spectra. This can be most easily seen for atomic levels, the wavefunctions of which can be expressed in terms of a radial part and a part containing spherical harmonics $Y_{l,m}$. As the operator $\vec{e}\cdot\vec{r}$ can also be represented in terms of spherical harmonics (e.g. $Y_{1,0}$ for linearly polarized light, or $Y_{1,\pm m}$ for circularly polarized light), the matrix element $\langle\psi_f|\,\vec{e}\cdot\vec{r}\,|\psi_i\rangle$ can be fully calculated by evaluating products of spherical harmonics. The particular mathematics of spherical harmonics allows the matrix element to be nonzero only for particular relations between l_f, l_i, m_f, m_i, which is the basis of the selection rules. Although being only strictly valid for atomic systems, this approach has also been successfully extended to approximately describe the behavior of electronic states at high symmetry points in solids. A more general treatment of dipole selection rules in solids has been developed on the basis of group theory [1].

It is useful to recall that the photoexcitation is only part of the entire photoemission process. Once the electron has been excited into the upper level – which takes place on a timescale of 10^{-15} s – we call it a photoelectron. However, this highly energetic or "hot" photoelectron must find a way to leave the crystal, which is only possible if the excitation occurs into states above the vacuum level E_{vac}. The proper quantum mechanical treatment of the photoemission process is the so-called *one-step model,* which – at least in principle – permits a quantitative interpretation of photoemission spectra. However, a more intuitive access to the underlying physics is provided by the simpler *three-step model* which we will discuss in the following.

2.1 Three-Step Model of Photoemission

This model separates the photoemission of a single[2] electron into subsequent processes dealing with (i) the *photoexcitation*, (ii) the *transport* of the hot electron to the surface, and (iii) the *transmission* of the electron through the surface into the vacuum (Fig. 4). On a quantum me-

[2]This single-electron picture is convenient, because it allows in many cases a qualitative interpreation of photoemission spectra on the grounds of band structure calculations within the framework of density functional theory. It neglects, however, electronic correlations in the electronic structure which can be significant in certain materials or materials classes.

chanical level these steps have to be connected in a suitable way in order to allow the electronic wave function to propagate from one step to the next.

2.1.1 Photoexcitation

We have already mentioned above that the central aspect in the photoexcitation step concerns dipole selection rules. These rules predict allowed electronic transitions based on the symmetry of the electronic wave functions involved. In an atomic picture, these selection rules take the form:

$$\Delta L = \pm 1 \tag{7}$$

$$\Delta m_L = 0, \pm 1 \tag{8}$$

The photon carries an amount of angular momentum of $|L| = 1$ with it's polarization state being determined by $m_J = 0$ (linear polarization) and $m_J = \pm 1$ (right- and left-hand circular polarization, respectively). Linearly polarized light can be represented as a superposition of right- and left-hand circularly polarized waves. To illustrate the action of these selection rules we take the example of the excitation of an atomic $2p$ level. According to Eq. (11) we will find two types of allowed transitions, which may contribute to the photoemission spectrum

$$p \rightarrow \begin{cases} d \text{ for } \Delta L = +1 \\ s \text{ for } \Delta L = -1 \end{cases} \tag{9}$$

For the evaluation of Eq. (8) it is useful to consider that atomic states are usually subject to spin-orbit coupling, which leads to a characteristic splitting of the atomic levels and leaves only the total angular momentum $J = L + S$ as a good quantum number. Consequently, our p-level splits into a $p_{3/2}$ and a $p_{1/2}$ state and with linearly polarized light, we will have allowed transitions of the type

$$\begin{array}{lcl} p_{3/2} & \rightarrow & d_{3/2} \\ p_{1/2} & \rightarrow & s_{1/2} \\ p_{-1/2} & \rightarrow & s_{-1/2} \\ p_{-3/2} & \rightarrow & d_{-3/2} \end{array} \tag{10}$$

whereas circularly polarized light gives us

$$\begin{array}{lcl} p_{3/2} & \rightarrow & d_{5/2} \\ p_{1/2} & \rightarrow & d_{3/2} \\ p_{-1/2} & \rightarrow & s_{1/2} \\ p_{-3/2} & \rightarrow & s_{-1/2} \end{array} \text{ for } (\Delta m_L = +1) \quad \text{and} \quad \begin{array}{lcl} p_{3/2} & \rightarrow & s_{1/2} \\ p_{1/2} & \rightarrow & s_{-1/2} \\ p_{-1/2} & \rightarrow & d_{-3/2} \\ p_{-3/2} & \rightarrow & d_{-5/2} \end{array} \text{ for } (\Delta m_L = -1) \tag{11}$$

Note that the dipole operator of the light acts only on the orbital part of the electronic wave function, i.e. on the spatial symmetries, but it cannot interact with the electron spin S directly. However, because spin-orbit coupling ties the spin to specific orbitals, a selective excitation can yield spin polarized photoelectrons even from nonmagnetic materials. This phenomenon is called optical spin-orientation [14] and is also the basis of all magnetodichroic effects observed in photoabsorption and photoemission [15]. In order to see how this works let us have

a closer look at the $p \rightarrow s$ transitions described by Eq. (11) (selection rules for crystalline symmetries see Appendix). The states $s_{1/2} \equiv |\uparrow\rangle$ and $s_{-1/2} \equiv |\downarrow\rangle$ may be regarded as pure spin states. For positive light circularity we have transitions starting at $p_{-3/2}$ and $p_{-1/2}$. Usually, the probabilities for the two transitions - which can be simply calculated from the Clebsch-Gordon coefficients [16] - differ by a factor of 3, i.e. the amount of photoelectrons excited into $s_{-1/2}$ is 3 times that for the $s_{1/2}$ state. If we assume that we have a nonmagnetic situation, the $s_{-1/2}$ and $s_{1/2}$ states will be energetically degenerate and a summation over the two photocurrent contributions will yield a spin polarization of the photoelectrons of $P = -50\%$. The same treatment for negative light circularity yields $P = 50\%$, i.e. a reversal of the circularity also reverses the sign of the photoelectron spin polarization. This optical spin-orientation effect in the $p \rightarrow s$ transitions is particularly exploited in spin-polarized GaAs photocathodes [17], but it can also be observed as a general photoemission phenomenon in basically all materials.
The quantity measured in the photoemission experiment is a photocurrent $I(h\nu)$, which is composed by transitions between all possible initial (i) and final states (f)

$$I(h\nu) \sim \sum_{i,f} |\langle f| \mathcal{O} |i\rangle|^2 \, \delta(\epsilon_f - \epsilon_i - h\nu) \tag{12}$$

Note that in this single particle picture the photoemission spectrum is represented by a series of sharp lines, which is not what is observed in the experiment. The reason of this discrepancy is the many-electron nature of the photoemission process, which will be discussed in sect. 2.3.
Once the photoelectron has been excited into the upper state it will propagate through the solid with a kinetic energy of $E_{kin} = h\nu - E_B$ according to energy conservation. The propagation direction is determined by the electron momentum $\hbar\vec{k}$, i.e. the electron wave vector $\vec{k}$ inside the crystal, which in turn is determined by a wave vector conservation law $\vec{k}_f = \vec{k}_i + \vec{q} \pm \vec{G}$. The relation between initial and final state wave vectors $\vec{k}_f$, $\vec{k}_i$ and the photon momentum $\vec{q}$ is given modulo a reciprocal lattice vector $\vec{G}$. In most cases, we can therefore focus our considerations on the photoemission spectra to the first Brillouin zone. For low photon energies well below 1000 eV the photon momentum can be safely neglected and particularly with respect to electronic band structures one then may assume *vertical* transitions, directly connecting the initial and final state. For higher photon energies, however, the photon momentum may become a crucial quantity.

2.1.2 Propagation

Inelastic mean free path – The propagation of the hot electron is captured in the second step of the three-step model. It takes into account that due to the strong Coulomb interaction in a solid the hot electron will suffer very efficient elastic and inelastic scattering processes, which affect both the energy and angular distribution of the photoemission spectrum observed outside the crystal. The main mechanisms are scattering due to electron-electron interactions and at defects. In particular, the inelastic scattering processes lead to a relaxation of the photoelectron towards the Fermi level. The effect of the inelastic scattering is usually discussed in the framework of an exponential damping of the photoelectron intensity along the path l

$$I(l) = I_0 \exp(-\frac{l}{\lambda_{in}}) \tag{13}$$

and is generally taken into account by the quantity λ_{in}, the *inelastic mean free path* (IMFP). The concept of the IMFP is essential to describe the finite information depth in a photoemis-

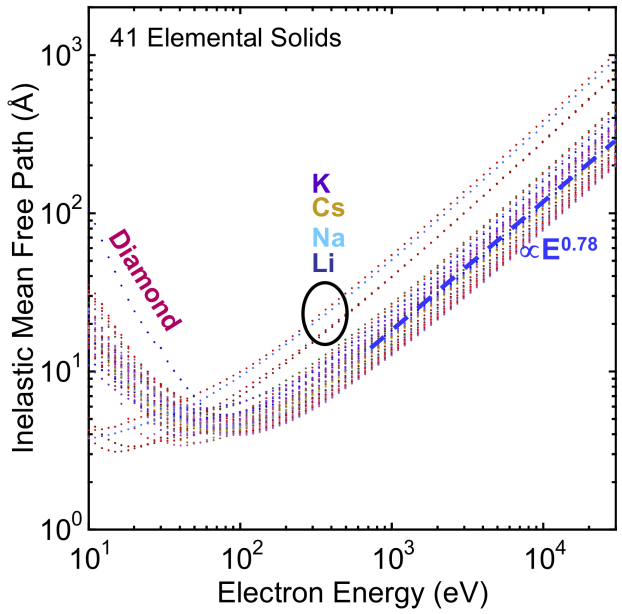


Figure 5: *IMFP values for 41 elements, calculated using the TPP-2M formula: Li, Be, three forms of carbon (graphite, diamond, glassy C), Na, Mg, Al, Si, K, Sc, Ti, V, Cr, Fe, Co, Ni, Cu, Ge, Y, Nb, Mo, Ru, Rh, Pd, Ag, In, Sn, Cs, Gd, Tb, Dy, Hf, Ta, W, Re, Os, Ir, Pt, Au, and Bi. Five "outlier" elements (diamond and the alkali metals) are included to illustrate the influence of the electronic structure characteristics. The dashed straight line for higher energies represents a variation as* $\lambda_{in} \sim E_{kin}^{0.78}$*, and is a reasonable first approximation to the variation for all of the elements shown. From [20].*

sion experiment. λ_{in} describes the average distance between two subsequent inelastic scattering events. This quantity is generally believed to follow a very similar behavior in different materials, which leads us to the well-known *universal curve* of the energy dependence of λ_{in}. A closer look, however, reveals that the curve is universal with respect to the general shape only. In quantitative respects the curve depends on the electronic structure of the element or material in question (Fig. 5).
Common to the IMFP curves is a minimum of λ_{in} of only a few Ångstroms at kinetic energies of ~100 eV and an increase towards both lower and higher energies. On this basis we can understand the high surface sensitivity of photoelectron spectroscopy at intermediate energies, which is both a virtue and a limitation. It can be – at least partly – overcome by exciting the photoelectrons with hard x-ray photons, i.e. photon energies of 6 - 10 keV (see Chapter 4.5). From Fig. 5 we see that at a kinetic energy of 10 keV the IMFP increases up to 10 nm. This energy-dependent variation of the information depth forms the basis for hard x-ray photoemission.

Photoelectron diffraction – While moving through the crystal the photoelectron can also be elastically scattered, giving rise to photoelectron diffraction (PD). This phenomenon is often also referred to as x-ray photoelectron diffraction (XPD) due to the higher excitation energies that are often used. In XPD a core-level photoelectron scatters from the atoms neighboring the emission site, so as to produce an angular anisotropy in the out-going photoelectron intensity [18]. The qualitative effects expected for the simple case of emission from the bottom atom in the diatomic molecule are shown in Fig. 6(a), and a quantitative calculation for emission from the C 1s subshell in an isolated CO molecule at 500 eV kinetic energy is shown in Fig. 6(b). Electron-atom elastic scattering is typically peaked in the forward direction, with this effect becoming stronger (that is, having a stronger and narrower forward peak) as energy increases [18]. For the CO case in Fig. 6(b), the intensity in the forward direction is in fact enhanced relative to that expected without scattering (I_0 in the figure) by about three times. Thus, one expects in XPD curves both a forward scattering peak (some-times referred to as forward focussing) along

the interatomic direction, as well as higher-order diffraction interference effects that one can also consider to be holographic fringes. Back scattering is weaker as energy increases, but Fig. 6(b) shows that, even at 500 eV, there are still interference fringes in the backward direction. Such XPD effects are very useful to determine the local atomic arrangement around an emitter atom. The XPD signals can be interpreted and modelled using the ingredients shown in Fig. 6(c). The polarization $\hat{\varepsilon}$ of the light influences the directionality of the initial photoelectron wave, and for emission from an s-subshell, the outgoing unscattered wave φ_0 has an amplitude proportional to $\hat{\varepsilon} \cdot \hat{k}$, where $\hat{k}$ is a unit vector in the direction of the photoelectron wave vector, and the photoelectron deBroglie wavelength will be given by

$$\lambda_e = h/\left|\vec{p}\right| = 2\pi/\left|\vec{k}\right|, \quad \text{in convenient units}: \ \lambda_e[\text{Å}] = \sqrt{150.5/E_{kin}[\text{eV}]} \tag{14}$$

Thus, an electron with 150 eV kinectic energy has a wavelength of about 1 Å, and a 1500 eV electron of about 0.3 Å, and these numbers are comparable to atomic dimensions. The outgoing photoelectron will elastically scatter from neighboring atoms j to produce wave components φ_j, and this process is describable in first approximation by plane-wave scattering, or more accurately by spherical-wave scattering. This scattering can be incorporated into a scattering factor f_j, which is furthermore found to be strongly peaked in the forward direction for energies above about 500 eV, as noted previously. The photoelectron wave components will also be

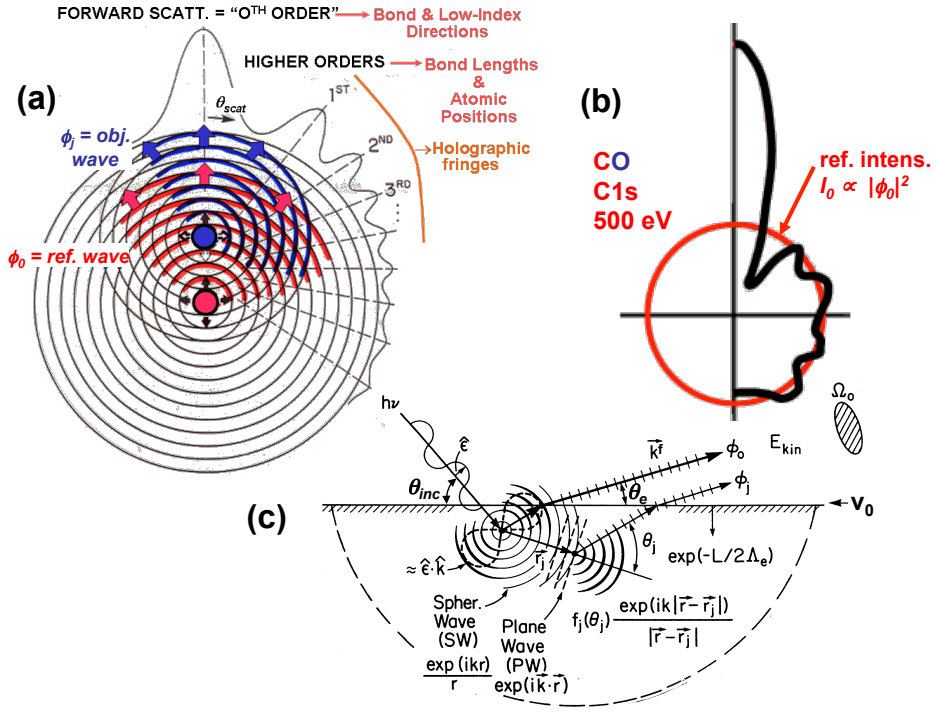


Figure 6: *Illustration of various aspects of photoelectron diffraction. (a) Simple diffraction features expected in emission from one atom in a diatomic system. (b) An accurately calculated diffraction pattern for C 1s emission from CO at a kinetic energy of 500 eV. Note the strong forward scattering peak, and other interference peaks or fringes extending from near the forward scattering direction to the backward scattering direction. (c) The basic theoretical ingredients required to describe photoelectron diffraction. From ref. [18], with calculations via ref. [19].*

inelastically attenuated as they traverses some total path length L in getting to the surface, with their amplitudes decaying as $exp(-L/2\Lambda_e)$. Finally, they will be refracted at the inner potential barrier V_0. Summing up all wave components (unscattered and scattered) and squaring then yields the diffraction pattern. Electrons can also be multiply scattered from several atoms in sequence, and accurate calculations of the resulting photoelectron diffraction patterns require including this for many cases, especially if scatterers are somehow lined up between the emitter and the detection direction, as along low-index directions in multilayer emission from a single crystal. Various programs are now available for calculating XPD patterns, with one web-based version being particularly accessible [19].

2.1.3 Transmission

In the final step of the three-step model the photoelectron has arrived at the surface and will leave the crystal. For this purpose, however, it has to pass through the surface potential barrier which matches the periodic potential inside the crystal to the vacuum outside. From simple quantum mechanics we know that an electron wave passing across a potential step Φ will be elastically scattered and diffracted, i.e. it will change its trajectory. In reality the surface potential barrier is not a step function, but smoothly varies as $\Phi(z)$. The details of the scattering process depend on the shape of $\Phi(z)$. The diffraction of the photoelectron in the surface potential is the reason that the wavevectors of the electrons inside the crystal $\vec{k}_{in}$ and outside in the vacuum $\vec{k}_{out}$ are not conserved. A conservation law exists only for the component parallel to the surface plane

$$k_{\parallel,in} = k_{\parallel,out} + G_{\parallel} \tag{15}$$

with $G_{\parallel}$ being a two-dimensional reciprocal lattice vector in the surface Brillouin zone. In order to relate the perpendicular component of the photoelectron wave vector $k_{\perp,out}$ to the electronic states inside the solid, more sophisticated methods for the band mapping have to be used [1]. The simplest one assumes the final electronic states to be described by a nearly-free electron parabola $E_f = \hbar^2 k_f^2/2m^*$ with an effective electron mass m^*. As the initial and final states are connected by vertical transitions in the band structure, k_f can be basically determined from E_f. This procedure fails, however, if the final state band structure deviates considerably from the nearly-free electron picture, which is the case particularly at hybridization regions. In this case, it is more reasonable to use final states from a band structure calculation for comparison. However, all these analyses will only yield a qualitative interpretation of the photoemission spectra. For a quantitative interpretation, a full photoemission calculation within a one-step model is needed.

2.2 One-Step Model of Photoemission

A considerable drawback of the three-step model is its limitation to a qualitative description of the photoelectron spectra. For a more quantitative description, first of all, the transition probabilities for all electronic excitations must be calculated on the basis of a realistic band structure for a semi-infinite system, for example, derived from density functional theory. The formalism must also take into account surface states or transitions into evanescent final states. Secondly, we must properly consider the multiple scattering events which the hot electron suffers in the final state. These are caused by the strong electron-electron interaction. This situation is more adequately described by the so-called one-step model, which considers the excitation and subsequent transport in a common framework. In particular, the multiple scattering of the

electron waves in the surface-near region and in the surface potential is treated in analogy to formalisms developed for low energy electron diffraction (LEED). In LEED an electron wave enters the crystal and subsequently goes through a multiple scattering process. If we invert the order on the time scale, we retrieve our final state in the photoemission process, with the excited electron propagating towards the surface. This state is therefore also called a *time-reversed LEED state* [21]. In terms of kinetic energy there is a gradual transition from the LEED to the photoelectron diffraction regime. PED effects can therefore be included into one-step photoemission theories.
On the basis of one-step photoemission theories one arrives at a reasonable quantitative description of the photoelectron spectra. This may even include effects due to spin-orbit coupling and the electron spin, in which case a relativistic Dirac-type formalism is involved. In this way it becomes possible to calculate magnetic dichroism and spin polarization spectra.

2.3 Refinement of the One-Electron Model

2.3.1 Electronic correlations

In the above discussions we have always implicitely assumed that the electronic system under investigation can be modelled within an effective single-electron picture. This has the advantage that the features appearing in the photoemission spectra may be directly related to specific interband transitions in the electronic structure, involving band states or core levels. As we have seen in Part A of this book, however, the single-particle picture may fail to capture the essential physics of a system, because it may underestimate the correlations in a many-electron system. This is true for the entire family of so-called highly-correlated systems, which includes transition metal oxides and other materials exhibiting phenomena such as high-temperature superconductivity, colossal magnetoresistance or multiferroicity. Such systems are usually described by theoretical approaches beyond simple LDA, for example, LDA+U or dynamical mean field theory (DMFT). The interpretation of photoemission results from such systems is more involved and must take into account the influence of the electron-electron interactions in all steps of the photoemission process.

2.3.2 Spectral shape of photoemission lines

There is, however, a second way through which electronic interactions enter the photoemission experiment. According to Fermi's golden rule (eq. 3) the δ-function taking care of the energy conservation implies all photoemission signatures to be infinitely sharp lines. This should hold particularly for core level photoemission lines. Inspection of the schematical picture in Fig. 3 already reveals that the photoemission line will have a finite width. This is due to the multielectron character of the photoemission process itself.
Whenever a photoelectron is excited to the upper level, it leaves behind a hole in the lower level. Strictly speaking the photoemission process converts an N-electron system into an $(N-1)$-electron system, if the photoelectron has left the crystal, before the hole has been filled again. The photoelectron and the hole interact with each other through Coulomb interaction, which may lead to a renormalization of the binding energies, the appearance of spectral satellites, and a finite linewidth of the spectral line. This has two profound consequences. First, the terms ground and excited state become a different meaning, because they rather refer to an N and $(N-1)$ electronic system, respectively. Second, photoelectron spectroscopy always measures an excited state of matter rather than the electronic ground state. In a somewhat larger picture

this is a nice illustration of one of the paradigms in quantum physics, according to which a measurement always affects and alters the system measured. Fortunately, modern condensed matter theory is able nowadays to handle many-electron systems both in the ground and excited state and can therefore provide a full photoemission calculation.
The role of electronic interactions in the photoexcitation spectrum is often taken into account within a Green function formalism [2]. The Green function $G(\vec{k},\epsilon)$ describes the behavior of a quasiparticle, which is "dressed" by the electronic correlations and the electron-hole interaction. Their influence is globally expressed by means of the complex self-energy Σ. Without going through the details of the formalism, from the Green function one can finally calculate the *spectral density function* $A(\vec{k},\epsilon)$, which may be compared to experimental results

$$A(\vec{k},\epsilon) = -\frac{1}{\pi}\mathrm{Im}G(\vec{k},\epsilon) = -\frac{1}{\pi}\frac{\mathrm{Im}\Sigma(\vec{k},\epsilon)}{[\epsilon-\epsilon_k-\mathrm{Re}\Sigma(\vec{k},\epsilon)]^2+[\mathrm{Im}\Sigma(\vec{k},\epsilon)]^2}. \tag{16}$$

A closer inspection of eq. 16 reveals that the real part of Σ introduces a renormalization of the energy ϵ_k of the spectral feature, whereas the imaginary part of Σ describes the finite lifetime of the quasiparticle state, resulting in a finite spectral width. As we will see below, the self-energy Σ can be conveniently employed to include further interactions, such as electron-phonon and electron-magnon coupling.
The spectral density function replaces the delta function in eq. 12. The total photocurrent is again determined by summing up over all dipole-allowed optical transitions between the many-electron states Φ_f and Φ_i weighted by the spectral density. We then arrive at the following description of the photocurrent [22]

$$\begin{aligned} I(h\nu) &\sim \sum_{f,i} |\langle\Phi_f|\,\mathcal{O}\,|\Phi_i\rangle|^2\, A_{ii}(\epsilon_f - h\nu) \\ &= \frac{1}{\pi}\sum_{f,i} |\langle\Phi_f|\,\mathcal{O}\,|\Phi_i\rangle|^2\, \frac{|\mathrm{Im}\Sigma(\epsilon_i)|}{[\epsilon_f-\epsilon_i-h\nu-\mathrm{Re}\Sigma(\epsilon_i)]^2+[\mathrm{Im}\Sigma(\epsilon_i)]^2}. \end{aligned} \tag{17}$$

This is the basis for modern photoemission calculations, which attempt a quantitative interpretation of the experimental data.

3 Techniques

3.1 Electron Spectrometers

During the last decade, we have seen a considerable improvement in photoelectron spectrometer technology, mainly driven by the continuous quest for improved spectral resolution. By now, commercially available energy analyzers may be able to achieve an energy resolution below 1 meV [23]. The most common type of photoelectron spectrometers nowadays are display-type energy filters which are able to efficiently acquire intensity distributions over a certain range of angles and energies within a single measurement. An example for such a hemispherical display spectrometer is given in Fig. 7. The photoelectrons moving away from the sample surface are accepted by a lens system, which defines the angular spread, i.e. the $k_\parallel$ value transmitted. The hemispherical capacitor employed to disperse the electrons according to their kinetic energy is usually operated at a fixed pass energy E_{pass} in order to keep the energy resolution constant

throughout a spectrum. The slit between the lens and the hemispheres separates the angular and energy information. The lens system therefore also has the task to accelerate/decelerate the electrons to the pass energy. After being dispersed in the electrostatic field a part of the electrons leaves the analyzer through a second aperture towards the areal electron detector. This usually comprises a combination of a multichannel plate (MCP) and position-sensitive read-out. The MCP consists of an array of narrow channels each being typical several 10 μm in diameter. In each channel a photoelectron is amplified by a factor of 10^4–10^6. This signal is then transported into the position-sensitive readout. The read-out may be a phosphor screen observed by an intelligent camera system which sorts and counts the events into a 3-D array in a computer.

Alternatively, there are resistive anode type detectors, which directly output voltage pulses to a multichannel analyzer. The specific design in Fig. 7 in fact features such a resistive anode arrangement called a delayline detector (DLD) [24] to make room for a second detector measuring the spin polarization of the photoelectrons (see below).

For specific purposes a wide variety of specialized electron spectrometers have been developed over the years. Most of them employ the electrostatic dispersion principle or a time-of-flight approach, in which the kinetic energy of the electrons is converted into a transit time along a defined trajectory. One of the challenges is to increase the angular acceptance of the analyzer in order to be able to capture a larger part of the photoelectron angular distribution in front of the sample. A very interesting design in this respect is the display-type spherical analyzer DIANA (Fig. 8) [25]. The electron trajectories emerging from the sample surface even at large emission angles are guided with high angular fidelity into the detector. The spectrometer is capable of

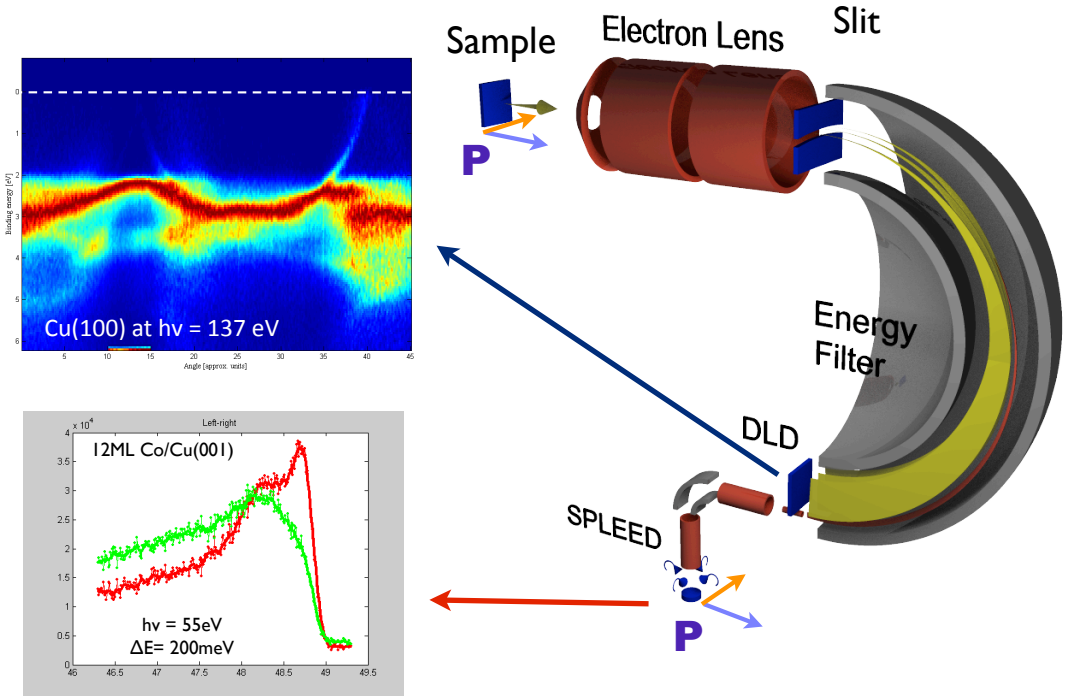


Figure 7: *Hemispherical photoelectron spectrometer with two-dimensional delayline detector (DLD) and SPLEED spin polarization analyzer operated at DELTA by the institute IFF-9.*

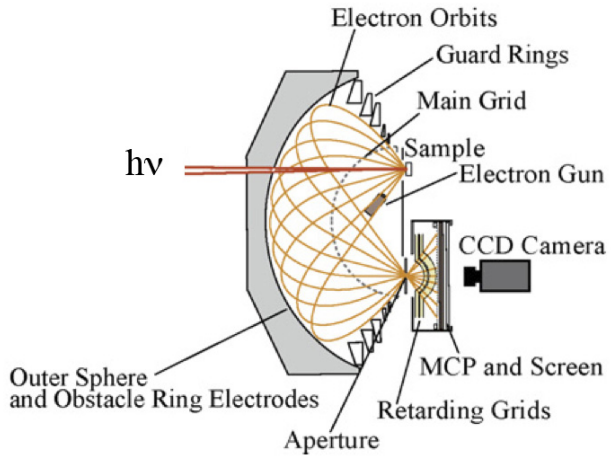


Figure 8: *Cross sectional scheme of the DIANA spherical capacitor spectrometer. The angular distribution of the photoelectrons is imaged onto a two-dimensional detector and captured by a CCD camera. From [25].*

mapping almost the entire half-space in front of the sample. This is particulary useful, for example, for photoelectron diffraction studies.

3.2 Spin Analysis

A "complete" photoemission experiment must analyze the photoelectron current with respect to all quantum numbers $|n, \ell, \vec{k}, s\rangle$. This includes also the electron spin, which carries important information about spin-dependent excitation and scattering processes in the solid. It can be shown that for an ensemble of electrons, the quantity spin can be expressed by a vector in real space, the spin polarization $\vec{P}$, the direction of which is defined by a quantization axis in the solid or the entire experiment [26]. Spin-dependent effects arise either through spin-orbit coupling or exchange interaction, the latter being a characteristic quantity in magnetic systems. Several types of spin polarization analyzers have been developed over the years . Their common principle of operation is based on the spin-dependent scattering of the photoelectrons off a target. The spin-dependence in the scattering process comes about by the same spin-dependent interactions mentioned above. In a simple picture, these interactions define a spin quantization axis and cause electrons with spin-up and spin-down to scatter with different probability into a direction perpendicular to this quantization axis. A counting detector placed in this direction will thus count different rates of scattered electrons for incident spin-up or spin-down photoelectrons, for example, $I^{\uparrow}(E)$ and $I^{\downarrow}(E)$. By subsequently orienting the spin-sensitive axis of the detector along the x, y, and z-axis, we can determine all three components of the spin polarization vector $\vec{P}(E)$.

There is only one spin polarization analyzer so far which is based on the exchange interaction. It involves low energy scattering ($E_s = 7$ eV) at a single-domain Fe(001) surface. Detectors exploiting spin-orbit coupling either involve high-energy Mott scattering at the atomic potential (several 10 keV up to 100 keV) or low energy scattering at the periodic potential of a solid (typically 100eV). Since the strength of the spin-orbit coupling increases with the atomic number *Z*, all spin-orbit scattering targets comprise heavy atoms, such as Au, W, or U.

In order to see how a spin detector is interfaced with the electron spectrometer, we choose the SPLEED detector as an example (Fig. 9). In this detector one effectively performs a spin-polarized low-energy electron diffraction experiment [27]. The incoming electrons hit a W(001) surface at normal incidence with a scattering energy of about 104 eV. The diffracted beams cre-

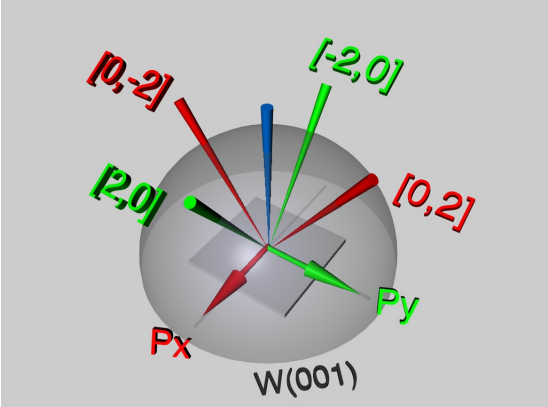


Figure 9: *Sketch of the SPLEED spin detection principle using a W(001) crystal. The spin polarization components P_x (red) and P_y (green) are measured simultaneously.*

ate a four-fold symmetric LEED diffraction pattern above the surface. Of particular importance are the {20} diffraction beams, because they provide the highest spin sensitivity at these scattering conditions. Because of symmetry reasons the SPLEED detector is sensitive to two orthogonal components of the spin polarization vector. The components P_x and P_y are determined from the intensity of the LEED reflexes by

$$\begin{aligned} P_x &= \frac{1}{S} \cdot \frac{I_{[0,2]} - I_{[0,-2]}}{I_{[0,2]} + I_{[0,-2]}} \\ P_y &= \frac{1}{S} \cdot \frac{I_{[2,0]} - I_{[-2,0]}}{I_{[2,0]} + I_{[-2,0]}} \end{aligned} \tag{18}$$

with the spin sensitivity S. This procedure is repeated for every data point of the spectrum and yields a spin polarization spectrum $P_{x,y}(\vec{k}, E)$, which can be used to calculate the spin-up and spin-down contributions of one vector component to the photoemission spectrum according to

$$I^{\uparrow}(E) = \frac{I_0}{2}(1+ \quad P(E)) \quad \text{and} \quad I^{\downarrow}(E) = \frac{I_0}{2}(1 - P(E)) \tag{19}$$

with the spin-averaged total intensity I_0.

3.3 Spectromicroscopy

The instruments discussed so far are dedicated to very good energy and angular resolution, but they usually do not provide a significant lateral resolution. As a consequence, the studies are restricted to relatively large (mm) and homogeneous samples. However, the progress in nanotechnology leading to ever smaller structures and also the need to study inhomogeneous samples on small length scales is the driving force for a new class of instruments, so-called spectromicroscopes.

Basically, there are two different avenues, which are currently followed in parallel. In the scanning approach one uses a conventional electron spectrometer, but tries to focus down the light beam by means of zone plate optics or capillary optics. Depending on the wavelength of the light and the quality of the optical components, a minimum spot size of less than 50 nm seems to be within reach. This light spot is then raster scanned with respect to the sample, recording a spectrum at each location [28].

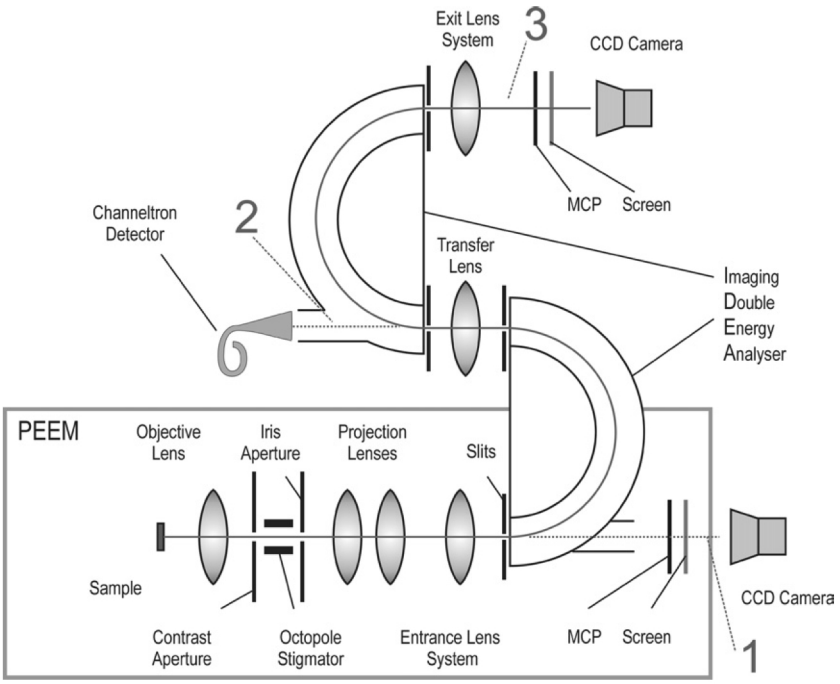


Figure 10: *Principle layout of the NanoESCA imaging spectrometer. From [29].*

In the full-field approach, the sample is illuminated by a more extended beam, and an electron-optical immersion lens microscope is used to image the surface in real time. This image is then passed through an energy analyzer, for example, an electrostatic one similar to the one discussed above. Behind the energy analyzer the now energy-filtered image is magnified onto the image detector, usually a MCP/phosphor screen unit with a slow-scan CCD camera. With these instruments, energy-filtered imaging with a resolution below 100 nm has already been shown. In Fig. 10 we show a different design which is employed in the NanoESCA instrument [29]. The energy filtering involves a system of two hemispherical analyzers connected by a transfer lens. The advantage of this layout is that imaging errors introduced by the first hemisphere are compensated by the second one, thereby improving the overall image quality. The instrument has three different imaging modes. The first one (1) takes the image from the PEEM column directly to the image detector. This corresponds to the normal unfiltered PEEM. The second operation mode (2) takes part of the electrons after the first hemisphere into a counting detector. The sample surface area where these electrons originate from is defined by means of an iris aperture and can be in the micrometer range. This mode is very interesting to measure small-spot spectra. Operation mode (3) corresponds to the energy-filtered imaging mode, i.e. the image is formed by electrons within a narrow energy window which can be adjusted, for example, to a particular core level photoemission line.

4 Selected Examples

In the following, we will discuss several examples illustrating different applications of photoemission spectroscopy covering band states and core levels.

4.1 Electronic and Chemical States

4.1.1 Valence state photoemission

The first example illustrates the mapping of the valence electronic states in a noble metal. A hemispherical display-type spectrometer like the one shown in Fig. 7 records an entire two-dimensional slice of the photoelectron distribution $E(\vec{k_f})$ in front of the sample in a single measurement. For a Ag(100) surface, which is illuminated by photons with energy $h\nu = 35$ eV, such a slice is displayed as a colour-coded map in Fig. 11 for photoelectrons emitted around an angle $\theta = 20°$ with respect to the surface normal. The energy scale is renormalized to the Fermi energy E_F, and the photoelectron intensity is represented as a function of binding energy E_B and emission angle θ.
For a more detailed analysis of the spectral features one may take cuts through the $I(E_B, \theta)$ distribution, resulting in different types of spectra. A cut at fixed binding energy yields $I(\theta)$, which is sometimes called a *momentum distribution curve* (MDC). A cut at fixed angle yields $I(E_B)$, which is called an *energy distribution curve* (EDC) and corresponds to a "classical" photoemission spectrum.
Although the distribution in Fig. 11 somewhat resembles a band structure, it is important to note that the data are not a simple picture of the bands, because the energy and angular position of the intensity maxima is determined by the transition matrix elements and thus by the initial state and final state bands. Nevertheless, we can already clearly discern different types of spectral features with large and smaller dispersion. In fact, a comparison to bulk band structure calculations of silver along the [100] (Δ) direction reveals that the spectral structure that starts at the Fermi level and bends downwards to the right originates from a strongly dispersing band of symmetry Δ_1, which has a strong free-electron, *sp*-like character. The more localized 4*d*-like states in silver give rise to the strong almost horizontal lines at binding energies below 4eV.
The position of the maxima in the MCD's and EDC's can be used to precisely determine the initial electronic state in the band structure by different approaches [1]. Usually this procedure is repeated for different emission directions and surface orientations in order to map the position of the initial state bands throughout the entire Brillouin zone, or at least along high-symmetry directions. Fig. 12 displays such a result for copper, with the data points being obtained from

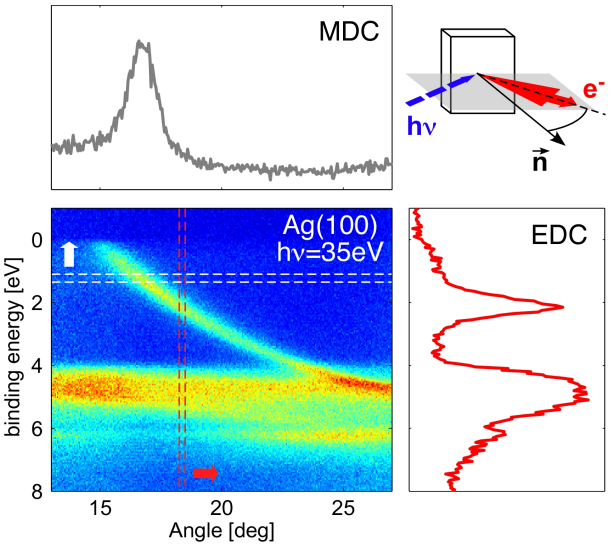


Figure 11: *Example of a two-dimensional $E(\theta)$ distribution recorded from an Ag(100) single crystal surface. The color code ranges from blue (no intensity) through yellow (medium intensity) to red (high intensity). Vertical and horizontal cuts through this distribution yield energy distribution curves (EDC) and momentum distribution curves (MDC), respectively. The broken lines bound areas of 5 lines on the detector which have been added up to the MCD (top) and EDC (right). Inset: Experimental geometry with the red triangle indicating the angular spread measured.*

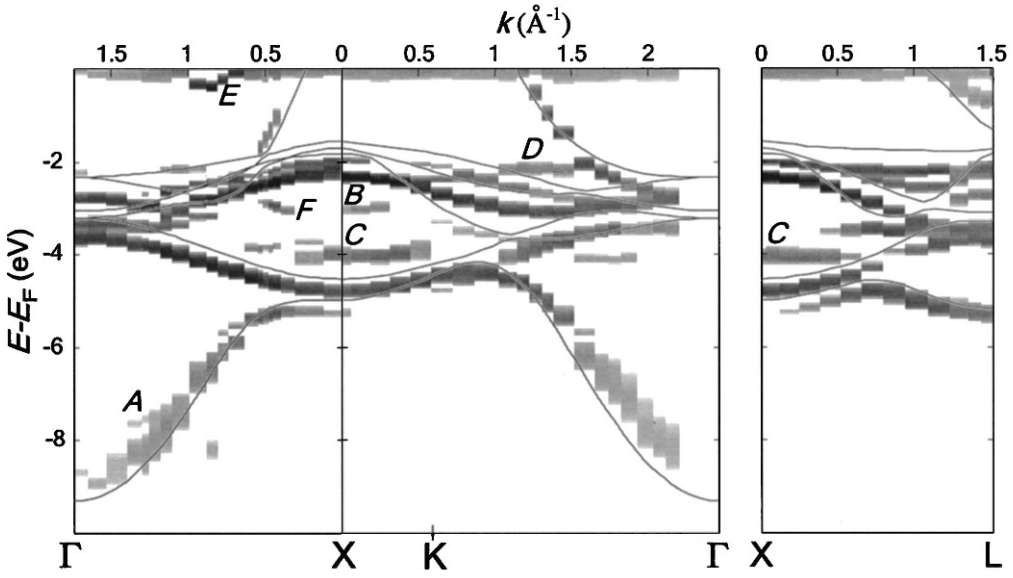


Figure 12: *Band mapping results for the bulk electronic states in a Cu single crystal along the [100]* $(\Gamma-\Delta-X)$*, ([110]* $(\Gamma-\Sigma-K)$*, and [111]* $(\Gamma-\Lambda-L)$ *directions. The bands are shifted from the DFT theoretical* $E(\vec{k})$*, shown by thin lines, due to excited-state self-energy effects. The constant line at* E_F *is due to the Fermi cutoff, and the peaks A–F are spurious structures due to multiple upper band composition, 1DOS maxima, and surface states. From [30].*

photoemission experiments and the lines representing a band structure calculation. The strongly dispersing bands starting at the Γ-point correspond to the free-electron like *sp*-type states. All other bands exhibit a weaker dispersion and have a strong *d*-type character, meaning that the electrons are more localized.

The band structure in Fig. 12 has been calculated within a relativistic scheme, i.e. it also contains the effects of spin-orbit interaction. Although copper is a material with low atomic number, spin-orbit coupling has been found to play an important role in the band symmetries, in particular, close to hybridization points. We also observe spectral signatures, which do not fit into the calculated bulk band structure ($A - F$) . A further analysis reveals that feature E is related to a surface state. The features B, C, and D are caused by transitions into regions with a strong one-dimensional density of states, where $k_\perp$ is not conserved. A hybridization of multiple upper bands leads to the appearance of feature A. Feature F is likely caused by surface state or surface resonance split off from the d bands [30]. We also note that the experimental data exhibit a systematic shift with respect to the calculated band. This is due to self-energy effects in the excited state.

4.1.2 Core level photoemission

The photoemission from the localized core levels gives rise to rather sharp spectral features at well-defined and characteristic binding energy values (cf. Fig. 3). These values are tabulated for the elements, for example, in the *X-Ray Data Booklet* [31], and range from several 10 eV for the shallow core levels up to 10 keV for the 1*s*-levels of heavy elements. Core level photoemission is often used to identify and quantify chemical species and is therefore also termed ESCA (Electron Spectroscopy for Chemical Analysis). For a sample consisting of one

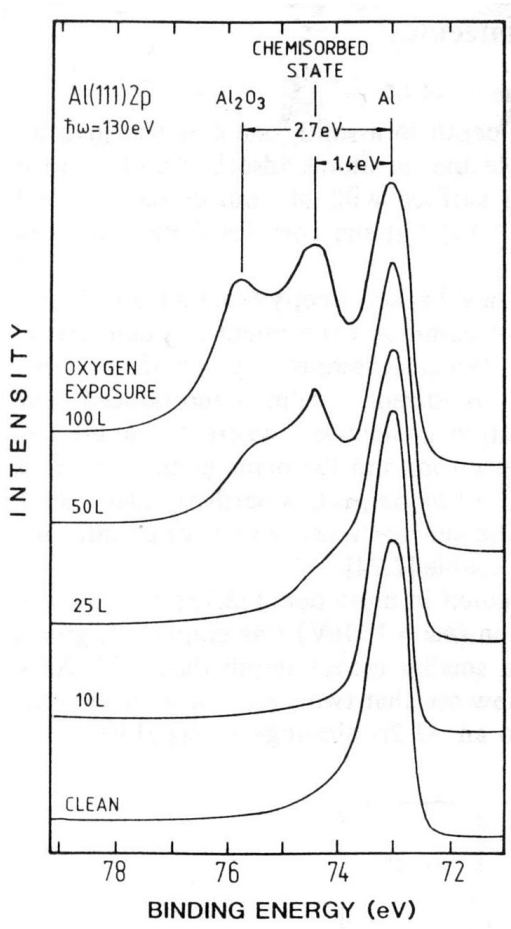


Figure 13: *Core level photoemission from Al(111). A surface reaction with oxygen leads to characteristic chemical shifts of the core level binding energies with respect to the clean surface. The amount of oxygen that the surface is exposed to is measured in units of Langmuir L ($1L = 10^{-6} mbar \cdot s$). From [1].*

chemical element only, the binding energy of the spectral feature is sufficient to unambiguously identify the element[3].

An example is given in Fig. 13, which compiles different spectra recorded for the Al 2*p* core level. The bottom most spectrum has been obtained from a clean Al(111) surface and shows the 2*p* core level peak located at a binding energy of $E_B = 73$ eV. A further inspection of the spectrum reveals that the spectral line has a finite width and an asymmetric shape. The line width is mainly determined by the lifetime of the core hole created in the excitation process and by the energy resolution of the electron spectrometer. The asymmetric line shape arises mainly due to the excitation of electron-hole pairs in the vicinity of the Fermi level. This corresponds to inelastic electron scattering of the photoelectron in the solid, effectively shifting some spectral weight to the low binding energy side of the peak. This asymmetric line shape may be modeled, for example, with the Doniach-Sunjic approach [32].

The binding energy of a given core level may change, as soon as we alter the chemical environment, for example, by a chemical reaction. Although the chemical bonds formed with an atom as a consequence of the reaction involve mainly the valence electrons, they may cause a charge transfer from or to that atom. This process modifies the electrostatic screening in the atom, ultimately resulting in a slight shift of the core level binding energy. These socalled *chemical* shifts form the basis of more elaborate ESCA approaches in determining the chemical composition

[3]Usually one measures several core level lines at different binding energies in order to increase the accuracy of the element analysis.

of complex alloys and compounds. An illustration for chemically induced core level binding energy shifts is given by the remaining spectra in Fig. 13. These spectra are obtained by exposing the Al(111) surface to different amounts of oxygen in successive steps. After an oxygen exposure of 25 L we start to see a weak spectral feature on the high binding energy side of the 2*p* level. After 50 L this feature has grown into a well-defined sharp peak $E_B = 74.4$ eV, which can be attributed to photoemission from Al surface atoms onto which oxygen has chemisorbed. In addition, a third peak starts to form at still higher binding energies. After dosing 100 L onto the Al(111) surface, this third signature at $E_B = 75.7$ eV has evolved into a clear peak, which can be attributed to photoemission from Al atoms bonded in an Al_2O_3 environment. We therefore see that the oxidation from metallic aluminium to alumina is accompanied by a chemical shift of the Al 2*p* core level by about $\Delta E_B = 2.7$ eV.

4.1.3 Fermi surface mapping

A particular aspect in modern photoemission spectroscopy is the so-called *Fermi surface mapping*. In order to see how this approach works, it is useful to recall that the data provided by the 2D display analyzers represent an intensity distribution $I(E, \vec{k}_\parallel)$. The electron wavevector $\vec{k}_\parallel$ parallel to the surface is defined by the experimental geometry. By varying the emission angles θ and ϕ (cf. Fig. 2) one obtains a set of slices through reciprocal space for different vectors $\vec{k}_\parallel = (k_x, k_y)$ in the surface plane. This data set can be condensed into a three-dimensional representation $E(k_x, k_y)$. An example for angle-resolved photoemission from a W(110) surface is shown in Fig. 14. The picture combines a vertical cut $I(E, k_x, k_y = 0)$ through the surface Brillouin zone (SBZ) with a horizontal cut $I(E = E_F, k_x, k_y)$. The intensity distribution $I(E, k_x, k_y = 0)$ reveals a clear dispersion of band segments along the high symmetry directions $\bar{\mathrm{S}} - \bar{\Gamma} - \bar{\mathrm{S}}$ in the SBZ, which can be compared to appropriate band structure calculations. The horizontal cut $I(E = E_F, k_x, k_y)$ depicts a two-dimensional map of the electronic states

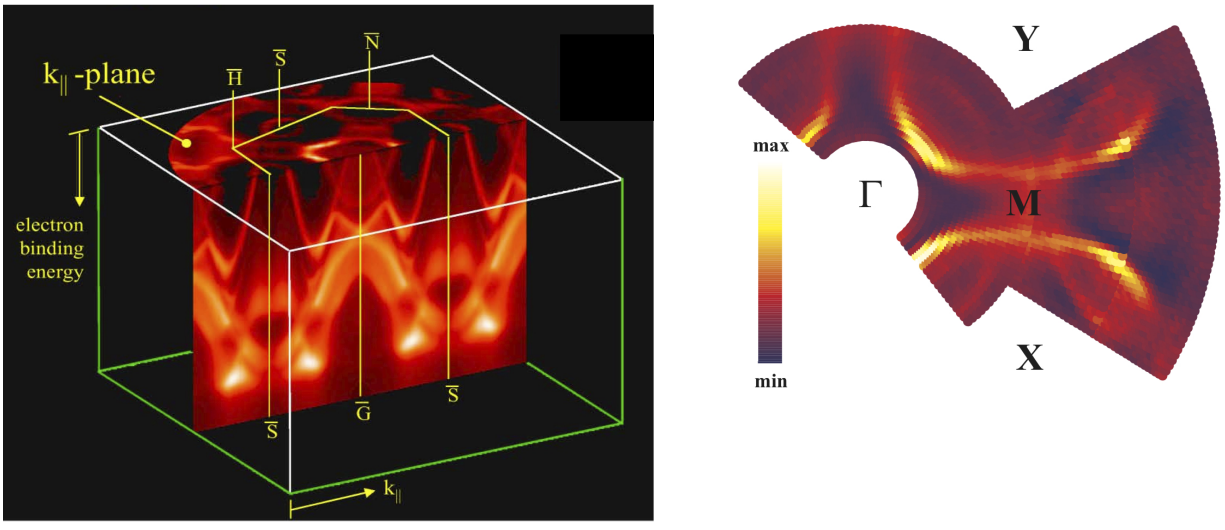


Figure 14: *(Right) Principle of Fermi surface mapping illustrated for photoemission from W(110). The plot compiles photoemission intensity distributions $I(E, \vec{k}_\parallel)$ for different emission angles, which can be stacked in a three-dimensional scheme. A cut through the stack at $E = E_F$ yields a two-dimensional map of the Fermi surface in the plane defined by $\vec{k}_\parallel$. From [33]. (Left) Fermi surface for Pb-doped Bi2212, i.e. $Bi_2Sr_2CaCu_2O_{8-\delta}$. From [34].*

at the Fermi energy and can thus be *related* to the Fermi surface. In the interpretation we have to keep in mind that the map in Fig. 14 contains matrix element and photoelectron diffraction effects. These have to be taken into account when comparing the data to theoretical predictions.

The details of the Fermi surface are crucial in determining the physical properties of materials, for example, the magnetic anisotropy in magnetic systems or the origin of superconductivity. In fact, the onset of superconductivity is accompanied by the formation of a small gap around the Fermi level. It is for this reason that Fermi surface mapping has become a standard tool in the investigation of high-T_C superconductors (HTSC). The example in Fig. 14 depicts the Fermi surface of Pb-doped $Bi_2Sr_2CaCu_2O_{8-\delta}$ (short form Bi2212) [34]. The data have been recorded in the normal state (T=120 K) with a He discharge source ($h\nu = 21.2$ eV). The main Fermi surface is hole-like and has the form of tubes (rings) centered around the X, Y high symmetry points. In addition, weaker intensity features are observed specifically around the M point. This so-called shadow Fermi surface is attributed to a spin-related origin [34]. A detailed understanding of the Fermi surface and their change with temperature are mandatory to understand the microscopic mechanisms leading to HTSC.

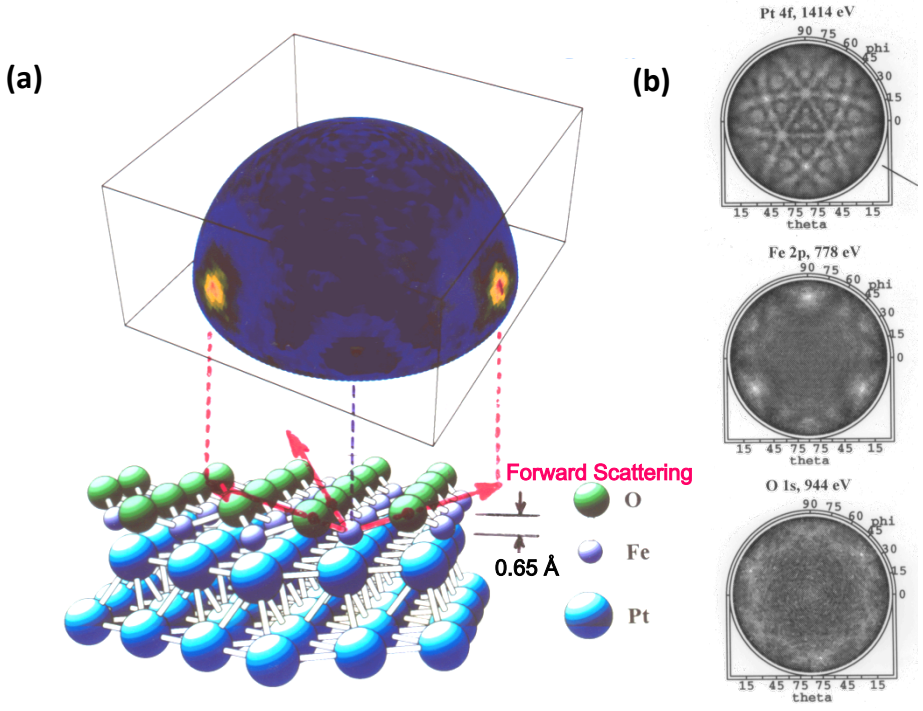


Figure 15: *X-ray photoelectron diffraction at 1486.7 eV excitation from a monolayer of FeO grown on Pt(111). (a) A full-hemisphere pattern for Fe $2p$ emission is shown, above the atomic geometry finally determined for this overlayer. (b) Diffraction patterns simultaneously accumulated for emission from Pt $4f$ (kinetic energy 1414 eV), Fe $2p$ (778 eV), and O $1s$ (944 eV). From Ref. [35].*

4.1.4 Photoelectron Diffraction

As one example of a photoelectron diffraction pattern, we show in Figure 15(a) the full hemisphere intensity distribution for Fe $2p$ emission at 778 eV ($\lambda_e = 0.44$ $\AA$) from a monolayer of FeO grown on a Pt(111) surface [35]. At this energy, the forward-peaked nature of fj is observed to create strong peaks in intensity along the Fe-O bond directions. The angle of these peaks can furthermore be used to estimate the distance between the Fe and O atoms in the overlayer, and it is found to be only about half that for similar bilayer planes in bulk FeO, as illustrated in the bottom of Figure 15(a). Figure 15(b) also illustrates the element-specific structural information available from XPD. The Pt $4f$ XPD pattern from the same sample is rich in structure due to the fact that emission arises from multiple depths into the crystal, with forward scattering producing peaks and other diffraction features along low-index directions. The Fe 2p pattern is here just a projection onto 2D of the 3D image in Figure 15(a). The O $1s$ pattern shows only very weak structure, as the O atoms are on top of the overlayer, with no forward scatterers above them, and only weaker back scattering contributing to the diffraction pattern. Comparing the Fe and O patterns thus immediately permits concluding that Fe is below O in the overlayer, rather than vice versa. Other examples of photoelectron diffraction in the study of clean surfaces, adsorbates, and nanostructure growth appear elsewhere [18, 36].

4.1.5 Microspectroscopy

Chemical Information – Laterally resolving photoemission techniques can yield a two-dimensional distribution of the electronic and chemical states at the surface. The current challenge in microspectroscopy is the improvement of the spectral and lateral resolution in order to increase the image contrast and reduce the minimum feature size detectable. The example in Fig. 16 has been obtained with a NanoESCA instrument (cf. Fig. 10) and represents the current status of the field. The sample is a GaAs/AlGaAs layer stack comprising layers of different thickness, which has been cut into a cross section through the layer stacking (schematic stacking sequence in the center of the image). The image has been acquired with Ga *3d* core level photoelectrons excited with photons of 400 eV energy.

The image contrast arises from the different relative Ga concentration in the GaAs and AlGaAs layers and directly reflects the layer stacking. The AlGaAs regions appear darker than the GaAs areas. The stack included a number of double layer structures, which appear in the imaged cross section as pairs of parallel lines. These can be employed to determine the lateral resolution in the experiment. The two pairs on the right hand side of the image are clearly separated. For the remaining structures on the left-hand side, a line scan analysis across the lines is necessary to extract further information. We can see that the line pair denoted as P3 in the image cannot be separated anymore into two individual features. If one assumes a 20%/80% criterion for the lateral resolution Δx, we find a value of Δx=273 nm.

On the left hand side of the image, we can discern a single line of darker contrast. This signal originates from a single AlGaAs layer (W3) of 38 nm thickness. The line scan reveals that this feature causes an image line with 270 nm full-width half maximum (FWHM). This is again the lateral resolution value that we have obtained above. This example demonstrates that features well below the resolution limit can still be made visible provided that the chemical contrast is strong enough.

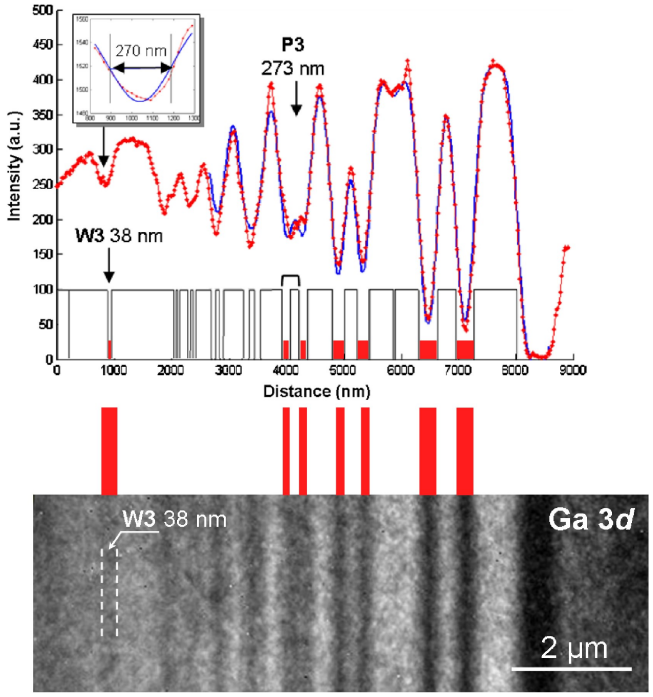


Figure 16: *Energy-filtered image from a GaAs/AlGaAs layered structure acquired at the maximum intensity of the Ga 3d core level peak (kinetic energy of 380.4 eV; incident photon beam of 400 eV). From [37].*

Fermiology – The electron-optical system of a spectromicroscope can be equipped with a transfer lens, by means of which the angular rather than the lateral distribution of the photoelectrons can be projected onto the image detector [38]. This approach permits a very efficient Fermi surface mapping, as an entire two-dimensional slice in reciprocal space can be recorded in a single measurement.

An example for such an experiment from an Ag(111) surface is given in Fig. 17. The picture represents a momentum map of the photoelectrons at the Fermi energy excited with photons of $h\nu = 40.8\,\mathrm{eV}$. It reveals the almost spherical Fermi surfaces in the first (center), 2$^{\mathrm{nd}}$ or next

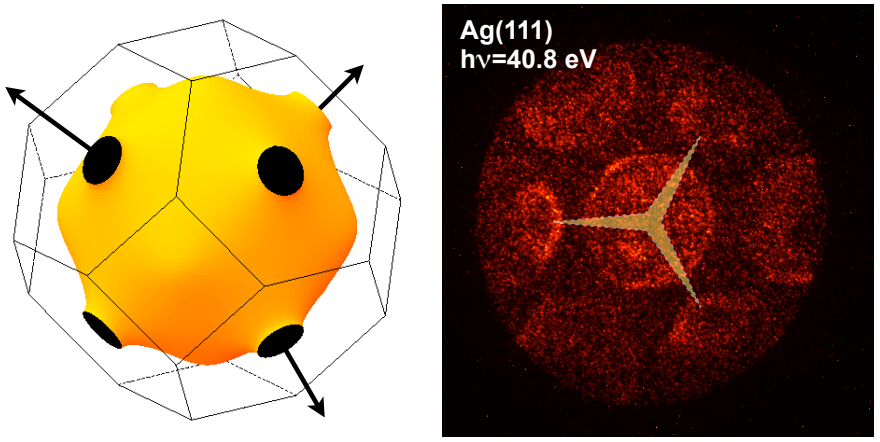


Figure 17: *(Right) Fermi surface mapping of an Ag single crystal along the [111] direction with the NanoESCA at $h\nu = 40.8\,\mathrm{eV}$ photon energy. The data show the Fermi surfaces in the nearest neighbor (in the direction of the star spokes) and the next nearest neighbor Brillouin zones. (Left) Calculated Fermi surface for Silver.*

neighbor (along the spokes of the star) and even the 3rd or next nearest neigbor Brillouin zone (BZ). The 1st and 2nd BZ are connected through the characteristic necks (schematic drawing in Fig. 17), whereas the 1st and 3rd BZ are separated by a larger gap. The necks, however, do not show up in the experimental momentum map. By switching between the real and reciprocal space imaging mode in the photoemission microscope, it is possible to directly correlate electronic and chemical structures.

4.2 Spin Effects in Photoemission

The electron spin can give rise to very peculiar phenomena in photoemission experiments. This is due to the fact that the electronic states are subject to two spin-dependent interactions: (i) spin-orbit coupling, and *(ii)* exchange interaction. Whereas spin-orbit coupling is mainly an atomic property, exchange-interaction is at the heart of the many electron system and is responsible for magnetic phenomena.

4.2.1 Ferromagnetic systems

A ferromagnet is characterized by a finite magnetization $\vec{M}$, i.e. a spontaneous long-range magnetic order below a critical temperature T_C. The magnetization is related to a lifting of the spin-degeneracy of the valence electronic states. As a consequence, the spin-up and spin-down bands are separated in binding energy by the exchange splitting $\Delta E_{exc}(\vec{k}, E)$. A spin-resolved photoemission experiment will therefore be able to directly distinguish between the spin-up and spin-down states, as the spin is preserved during the optical transition.

Fig. 18 gives an example for spin-resolved photoemission from the Fe(001) surface. The intensity spectrum (upper panel, solid line) exhibits a strong peak close to the Fermi level and a weaker spectral feature at around 3 eV binding energy. The sample is magnetized within the surface plane and the spin polarization distribution is measured along the magnetization direction (bottom panel). It reveals some structure in the region between E_F and 2 eV binding energy, and levels off to an almost constant value towards higher binding energies. From the intensity and spin polarization distributions we can calculate the spin-up and spin-down contributions to the photoemission spectrum according to eq. 19. These are represented by the closed and open symbols in the upper panel of Fig. 18. We can easily see that these partial spin spectra differ significantly from each other. In a qualitative picture the features marked by A-D can be related to transitions different initial state bands.

According to the nonrelativistic selection rules (see Appendix), the experimental set-up allows access to initial states of mainly Δ_1 and Δ_5 symmetry. By comparison to band structure calculations, the feature *B* can be ascribed to direct transitions from the $\Delta_5^\downarrow$-band lying closely below E_F. The other features appear in the majority spin spectrum: peak *A* is located at a binding energy of $E_B = -0.7 \pm 0.2$ eV and originates from an initial state of $\Delta_1^\uparrow$ symmetry, whereas feature C positioned at $E_B = -2.5 \pm 0.2$ eV results from the $\Delta_5^\uparrow$-band. Finally, a feature labeled D is located at $E_B = -3.4 \pm 0.2$ eV and originates from the strongly dispersive $\Delta_1^\downarrow$ band, which starts at the Γ-point. These findings are in good agreement with previous photoemission studies on Fe(001) as presented, for example, in Ref. [39].

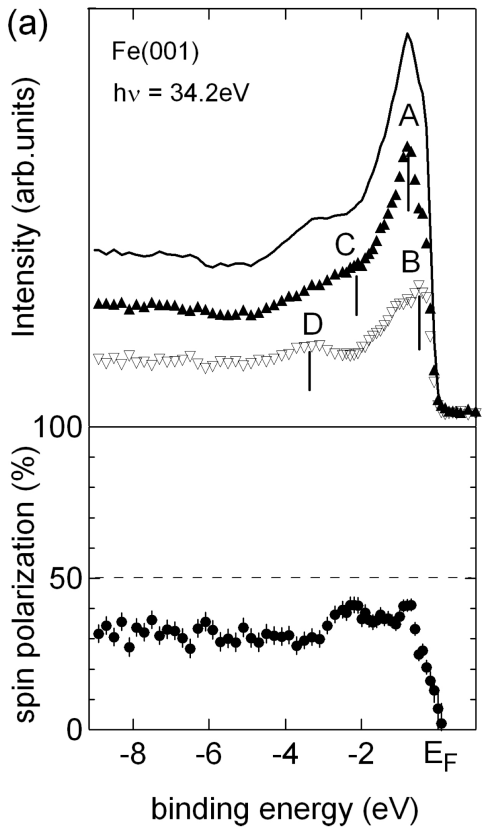


Figure 18: *Top panel: Spin-resolved photoemission spectra from Fe(001) along the surface normal at a photon energy of $h\nu = 34.2$ eV for p-polarized light impinging at 45° angle of incidence. The intensity spectrum (solid line) is decomposed into the spin-up (▲) and spin-down (▽) components. These have been extracted on the basis of the spin polarization distribution (bottom panel). The spectral features marked A-D are explained in the text.*

4.2.2 Optical spin orientation

Even in the absence of magnetic interactions, it is possible to observe spin-polarized photoelectrons and relate them to the symmetry of the electronic states. This phenomenon is called "optical spin orientation" and the microscopic mechanism is provided by spin-orbit coupling, as we have already discussed in sect. 2.1.1. The effects are large, if the spin-orbit coupling in the occupied states is strong.

As an example, Fig. 19 shows spin-resolved photoemission data for the W $4f$ shallow core levels obtained with linearly polarized light. The geometry was chosen such that the light impinges on the W(110) surface at a glancing angle of 17°. Symmetry arguments require that the spin-polarization vector is oriented perpendicular to the plane spanned by the direction of incidence an the surface normal [41]. These states show a clear spin-orbit splitting of about $\Delta E_{so} \approx 2.5$ eV between the $4f_{7/2}$ and $4f_{5/2}$. We see that the partial intensity spectra of spin-up (▲) and spin-down (▽) differ significantly at the peak positions, resulting in a positive spin polarization at the $4f_{7/2}$ emission line, whereas the $4f_{5/2}$ level exhibits a negative spin polarization (bottom panel). This spin polarization reversal between the spin-orbit split levels is an instrinsic feature of the optical spin orientation process, because the total spin polarization integrated over all spin-orbit split levels is required to vanish for symmetry reasons – at least in nonmagnetic materials. Note that for a given experimental geometry the sign of the spin polarization is unambiguously connected to the symmetry of the electron states involved in the optical transition. This assertion also holds for band states in a solid and allows a detailed analysis of spin-orbit effects in the band structure on the basis of spin-polarized photoemission experiments [42].

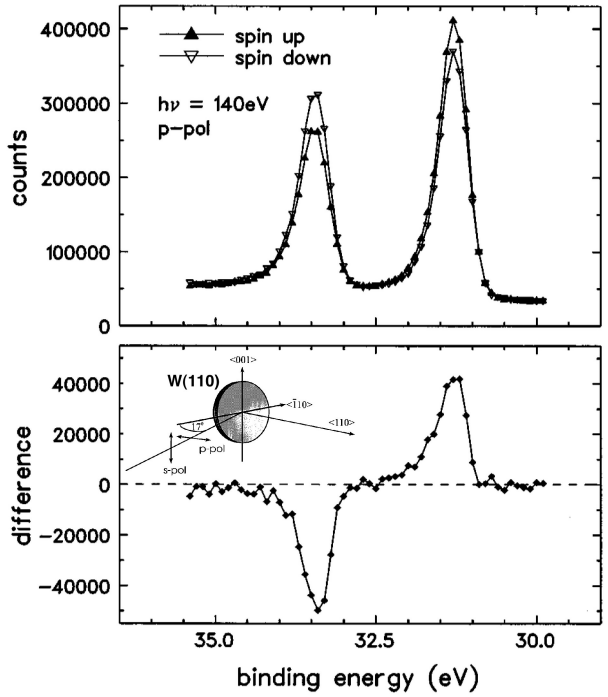


Figure 19: *Spin-resolved W 4f energy distribution curves (EDC's) measured with p-polarized light of hν=140 eV photon energy. The spin polarization vector is oriented perpendicular to the plane spanned by the direction of light incidence and the surface normal. The integral of the difference (lower panel) over the binding energy vanishes within 1% relative to the integral of the absolute value of the difference. From [40].*

4.2.3 Rashba States

The spin-polarization effect described in the previous section is due to the intraatomic spin-orbit interaction in the Hamiltonian (eq. 2). This term has a specific structure. In the field of spin-dependent transport there is a strong desire to control the electron spin in semiconductors by electric fields. In order to describe this situation in a planar configuration, one often uses the Rasbha-Bychkov Hamiltonian [43]. Interestingly, it has a very similar mathematical form

$$H_{RB} = \alpha(-i\vec{\nabla} \times \vec{E}) \cdot \vec{\sigma} \tag{20}$$

with the Rashba constant α and the effective electric field $\vec{E}$. One expects maximal effects of the Rashba Hamiltonian when the electric field, the electron momentum and the electron spin are mutually orthogonal.

In two-dimensional (2D) systems with broken inversion symmetry, this spin–orbit interaction causes spin separation of the moving electrons – which is why it is interesting for spintronics. However, the inversion symmetry of the potential is also naturally broken at any crystal surface or interface. As a consequence, electronic states localized at a surface/interface should be spin-split although this splitting can be quite small. In fact, the Rashba interaction at crystal surfaces becomes sizeable only when it couples to the large intra-atomic spin-orbit interaction. The gradient of the surface potential by itself is not sufficient to cause a directly observable splitting of the surface/interface electronic bands into spin subbands [44]. Therefore, this interaction plays an important role only if high-Z elements are involved at the surfaces or interfaces.

It is well-known that some noble metal surfaces exhibit pronounced surface states. This is also true for the unreconstructed Au(111) surface, which exhibits a Shockley-type surface state at the center of the surface Brillouin zone (SBZ), i.e. at the $\bar{\Gamma}$-point. This situation is depicted in Fig. 20. The surface state is characterized by a parabolic dispersion with $k_{||}$. Due to the Rashba

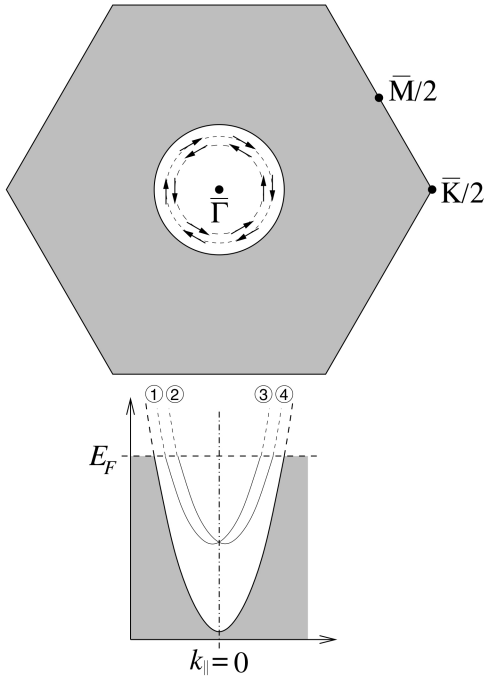


Figure 20: *Upper panel, section of the surface Brillouin zone of the unreconstructed Au(111) surface. The $\overline{\Gamma K}$ distance is $\pi\sqrt{32/3}a = 1.45\mathring{A}^{-1}$. Lower panel, schematic view of the split surface state dispersion in a cut through $\overline{\Gamma}$. From [45].*

interaction the surface state will spin-split, forming two concentric ring-shaped Fermi surfaces with opposite spin polarization in the SBZ.

Indeed this splitting can be clearly seen in a high-resolution photoemission experiment, (Fig. 21). The gray-scale intensity map represents a slice through the Brillouin zone along the direction $\bar{\Gamma}\bar{\mathrm{K}}$. Without the Rashba interaction there would be only one parabolic trace centered around $k_{||} = 0$, corresponding to the dispersion of the surface state. The Rashba interaction introduces a symmetric splitting resulting in two parabolic traces with opposite spin polarization. The energy and momentum distribution curves show that the two traces are only degenerate at $k_{||} = 0$, but separated otherwise.

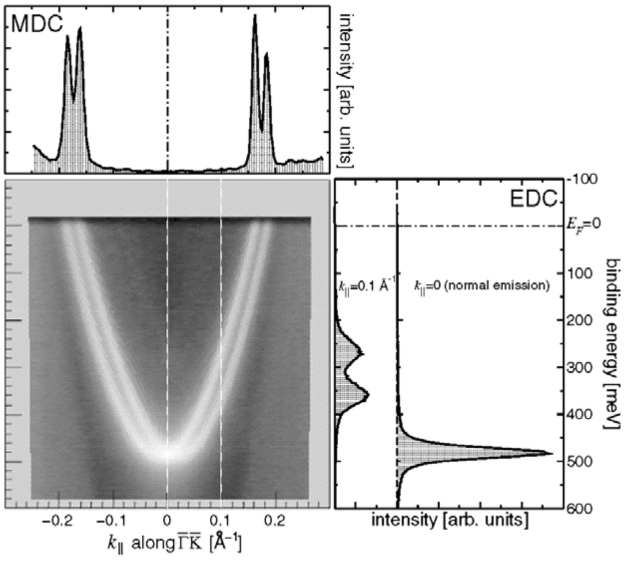


Figure 21: *Photoemission intensity of the Shockley state on Au(111) as a function of energy and momentum $I(E_B, k_{||})$ (white means high intensity). The top panel shows a cut at constant energy $E = E_F$ (MDC); the right-hand panel gives the energy distribution curves (EDCs) at $k_{||} = 0$ and $k_{||} = 0.1\,\mathring{A}^{-1}$. From [45].*

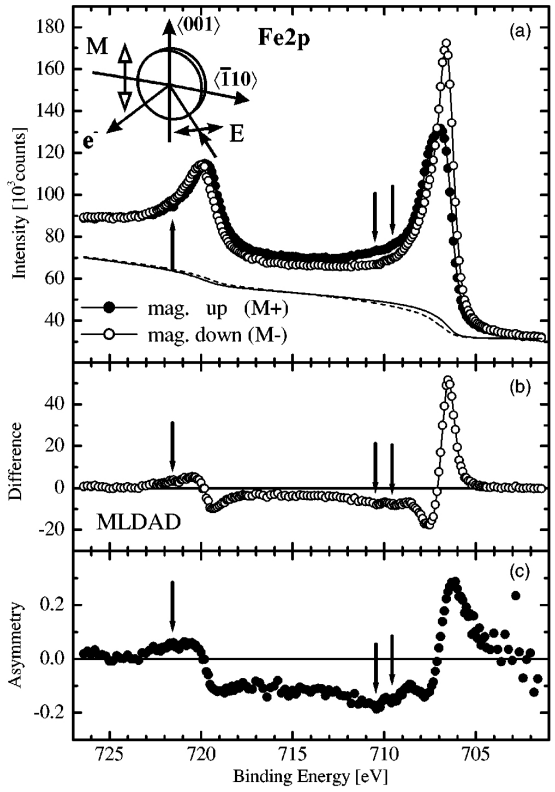


Figure 22: *(a) Fe 2p photoemission spectra and Shirley background of 15 ML Fe / W(110) excited with p-polarized radiation ($h\nu$ =850 eV) for magnetization up and down (M+, M−). The inset shows the experimental geometry. (b) The intensity difference (MLDAD) of the curves from (a). (c) MLDAD asymmetry (without background). The arrows mark the position of correlation-induced satellites. From [47].*

The (111) surfaces of silver and copper have very similar Shockley surface states, but a Rashba splitting could not be confirmed yet. This may be due to the smaller intraatomic spin-orbit coupling as mentioned above.

4.2.4 Magnetic Dichroism in Photoemission

What happens, if we have an experimental situation as described in sect. 4.2.2, but our sample is actually ferromagnetic? Let us take the example of a $2p$ core level. The ferromagnetic state is responsible for a spin-dependent energy splitting of the electronic states – not only in the valence states, but also in the core levels. These split according to their magnetic quantum number m_J, i.e. the $2p_{3/2}$ level splits into 4 sublevels ($m_{3/2}$, $m_{1/2}$, $m_{-1/2}$, $m_{-3/2}$), the $2p_{1/2}$ into two. The transition matrix elements depend on m_J *and* the orientation of the magnetization. As a consequence, the fine structure of the intensity spectrum depends on the magnetization direction. This phenomenon is called magnetic dichroism and is observed for both core levels and valence states [46].

This effect is shown in Fig. 22 for the $2p$ core level photoemission from Fe. Note that we have a very similar geometry as in experiment described in sect. 4.2.2. The magnetization vector is oriented perpendicular to the reaction plane. The upper panel compiles the photoemission spectra across the spin-orbit split $2p$ levels. We can clearly see that the spectra differ significantly for opposite magnetization directions. The difference of the two spectra is plotted in the center panel and reveals characteristic bipolar signatures at the position of the core levels. We also note that the polarity of these features reverses between the $2p_{3/2}$ and the $2p_{1/2}$. This is consistent with the spin polarization change in the optical spin orientation experiment in sect. 4.2.2. In fact, as a general rule, optical spin orientation phenomena in nonmagnetic materials

are taking the form of magnetic dichroisms in ferromagnets.
The magnetic dichroism signal is often expressed as an intensity asymmetry A

$$A = \frac{I(M+) - I(M-)}{I(M+) + I(M-)}. \tag{21}$$

which reveals a similar spectral dependence than the difference. Additional weak spectral features are related to correlation effects (see Sect. 4.3). As the experiment has been performed with linearly polarized light, the effect is also termed magnetic linear dichroism in the photoelectron angular distribution (MLDAD). The latter points out that the size and sign of the magnetic dichroism depends strongly on the emission angle of the photoelectrons analyzed. A closer theoretical analysis shows, that the MLDAD is actually an interference effect between the two photoemission channels into s and d final states [48].

4.3 Electronic Correlations

It is well established nowadays that photoemission spectra of narrow-band materials, such as the elements of the d transition-metal series and their compounds, cannot be entirely explained within a one-electron picture. This is due to the presence of local correlations between electrons in the partially filled d band. Experimental band mapping and its comparison with theoretical results can be a powerful tool to directly investigate correlation effects. It has to be realized, however, that the correlated electron picture is less transparent than the single particle model. The interactions due to the electronic correlations lead to a "dressing" of the single particle, i.e. when the particle moves in the solid it is always screened by these many-particle interactions. This system of particle and interaction cloud may be seen as a new quasiparticle. The respective many-electron calculations result in quasiparticle spectral functions rather than conventional band structures, which is a significant conceptual difference.
From all d transition metals, Ni has the narrowest bands and exhibits the strongest correlation effects. This can be seen in Fig. 23. Panel (a) reproduces a set of experimental angle-resolved

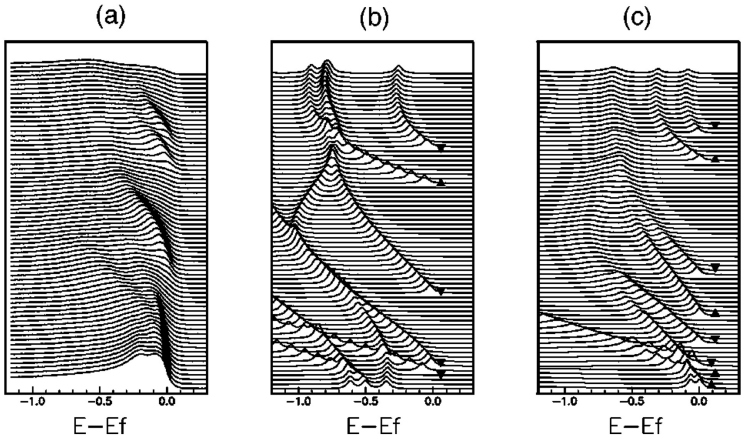


Figure 23: *Comparison between angle-resolved photoemission spectra from a Ni(110) surface at $h\nu = 21.2\,\mathrm{eV}$ (a), single particle local-density approximation (LDA) (b), and quasiparticle calculations results (c). The polar angle ranges from 0° (bottom) to 70° (top). The spin character is indicated by ▲ and ▼. From [49].*

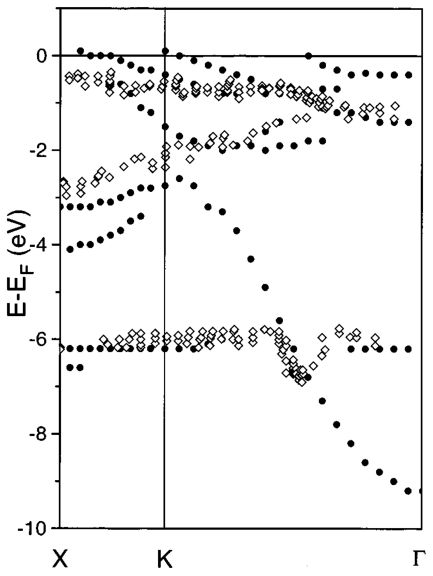


Figure 24: *Comparison between the calculated dispersion of quasiparticle states (•) for majority-spin bands and angle- resolved spin-integrated photoemission results (◊) of Ref. [51]. From [52].*

photoemission spectra, which have been recorded for different emission angles from normal emission up to 70°. It is obvious that the spectral features disperse with the emission angle. A comparison with a single particle particle calculation in the LDA approximation (panel b), however, predicts a much stronger dispersion of the bands than observed in the the experiment. In particular, strong spectral features should also be expected at binding energies larger than 0.5 eV. This is also not observed in the experiment. Furthermore, the exchange splitting between bands of the same symmetry is calculated about twice as large, as observed in spin-resolved experiments ($\Delta E_{exc} \simeq 300\,\mathrm{meV}$ [50]).

The quasiparticle spectral functions calculated within a multiorbital Hubbard model for the experimental geometries are compiled in panel (c) of Fig. 23. The inclusion of correlation effects strongly modifies the spectra: all the structures are pushed up towards E_F by self-energy corrections reproducing much more closely the experimental results both in terms of energy position and dispersion. The spin dependence of the self-energy, arising from the different efficiencies of the scattering channels involving majority- and minority-spin electrons, strongly affects the spin polarization of the quasiparticle states. For this particular region in k space, four spin-up and four spin-down bands are theoretically predicted in the energy region of interest. While in the single-particle picture one spin-up band and four spin-down bands cross the Fermi energy, all four spin-up bands come close to E_F after the inclusion of correlation effects. Moreover, the energy separation between the spin-up and spin-down bands between $\theta = 50°$ and 60° is reduced by self-energy corrections. All this is in excellent agreement with the experimental data.

In addition to a spin- and energy dependent renormalization of the quasiparticle states due to the self energy, the correlations also lead to the appearance of new spectral features, which are completely absent in the single particle band structures. The most prominent feature in Ni is the famous "6 eV satellite". This is depicted in Fig. 24, which shows the calculated dispersion of the majority spin quasiparticle states. These are compared to spin-integrated, angle-resolved photoemission results. In the region close to E_F we observe the spin-dependent energy renormalization already discussed above. In addition, we find a strong dispersing feature corresponding to the sp-type band. At about 6 eV below the Fermi level, however, there appears

a new non-dispersing feature. This is the correlation-induced satellite, which indeed turns out to be of majority-spin character in spin-resolved photoemission experiments [53].

4.4 Kinkology

The on-site Coulomb interactions leading to the correlation phenomena discussed above are relatively strong and thus lead to large effects in the band structure. High-resolution photoemission nowadays provides the opportunity to study also the influence of much weaker interactions affecting the electronic system, for example, electron-phonon or electron-magnon interactions. As the analysis procedure is connected close to finding and identifying kinks and precisely measuring the spectral width in the dispersion of the quasiparticle states, this field is sometimes coined "kinkology".

4.4.1 Electron-phonon interaction

The interaction of different quasiparticles, such as electrons and phonons, results in a crossing and hybridization of their respective dispersion relations. At the position in k-space where such crossings occur, the states involved are shifted in energy with respect to the noninteracting case. As phonons have very low energies of the order of 100 meV the respective modifications of the dispersion behavior of the electronic quasiparticle states due to the electron-phonon interaction will be confined to a narrow region below the Fermi level. Formally, the electron-phonon interaction can be considered as an additional contribution to the self energy Σ.
All characteristics of the electron-phonon coupling (EPC) are described by the *Eliashberg function* $E(\omega, \varepsilon, \vec{k}) = \alpha^2(\omega, \vec{k})F(\omega, \varepsilon, \vec{k})$, the total transition probability of a quasi-particle from/to the state $(\varepsilon, \vec{k})$ by coupling to phonon modes of frequency ω [54]. Information about the Eliashberg function can be obtained from the angle-resolved photoemission spectra, both through the EPC distortion of the quasi-particle bands near the Fermi energy and the temperature-dependent linewidth. If $\varepsilon_0(\vec{k})$ is the bare quasi-particle dispersion of a surface state without EPC, then the measured dispersion $\varepsilon(\vec{k})$ with electron-phonen coupling is given by

$$\varepsilon(\vec{k}) = \varepsilon_0(\vec{k}) + \mathrm{Re}\Sigma(\vec{k}, \varepsilon) \tag{22}$$

The screening of the electrons by the lattice is represented by the self-energy function $\Sigma(\vec{k}, \varepsilon)$. The imaginary part of the self-energy is related to the EPC contribution to the lifetime τ of the excited electronic states,

$$1/\tau = 2\mathrm{Im}\Sigma(\vec{k}, \varepsilon, T) \tag{23}$$

Based on these considerations the influence of the electron-phonon coupling has been investigated in Be($10\bar{1}0$) [55]. The photoemission data in Fig. 25 (left) show the dispersion of the surface states S1 and S2 as bright features. The experimental dispersion of the quasiparticle band $\varepsilon(\vec{k})$ (Fig. 25, right) is compared to the expected dispersion of the surface state without additional interactions $\varepsilon_0(\vec{k})$. This comparison reveals a weak, but distinct deviation of the experimental data from the parabolic dispersion close to the Fermi energy. This kink is the spectral signature of the electron-phonon coupling.

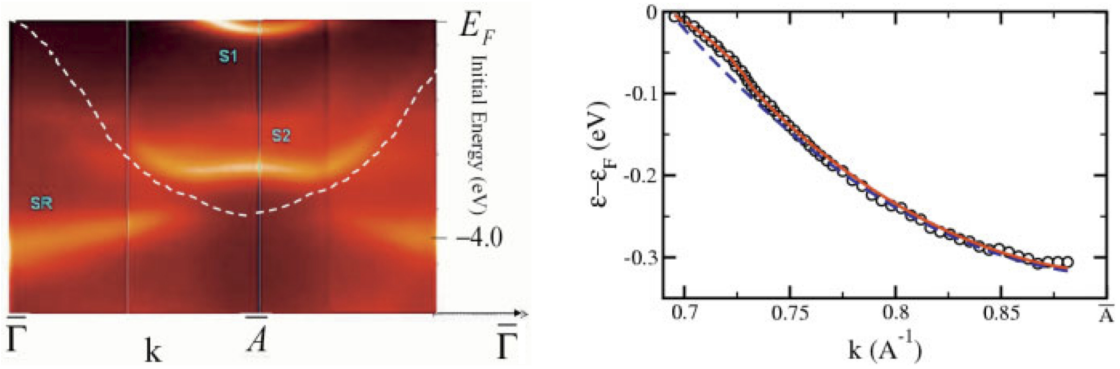


Figure 25: *(Left panel) Energy vs. momentum photoemission display of the two surface state bands S1 and S2 on Be($10\bar{1}0$). The dashed line is the bulk band edge. Data taken at 30 K at 40 eV photon energy. (Right panel) Quasi-particle dispersion determined from momentum distribution curves (circles) obtained at 24 eV photon energy. Dashed blue line is the bare particle dispersion $\varepsilon_0(k)$ and the red line is the fit to the data from the extracted Eliashberg function. From [55].*

4.4.2 Electron-magnon interaction

In a magnet we have collective excitations of the spin system – magnons. These quasiparticles have energies also in the 100 meV range. We should therefore expect that electron-magnon interaction leads to the appearance of kinks in the band structure of ferromagnetic materials. This is demonstrated for the photoemission from the Fe(110) surface. The ARPES data (Fig. 26) show the spectral distribution of the surface state photoemission close to E_F at the center of the surface Brillouin zone. A careful analysis of surface state dispersion reveals a characteristic

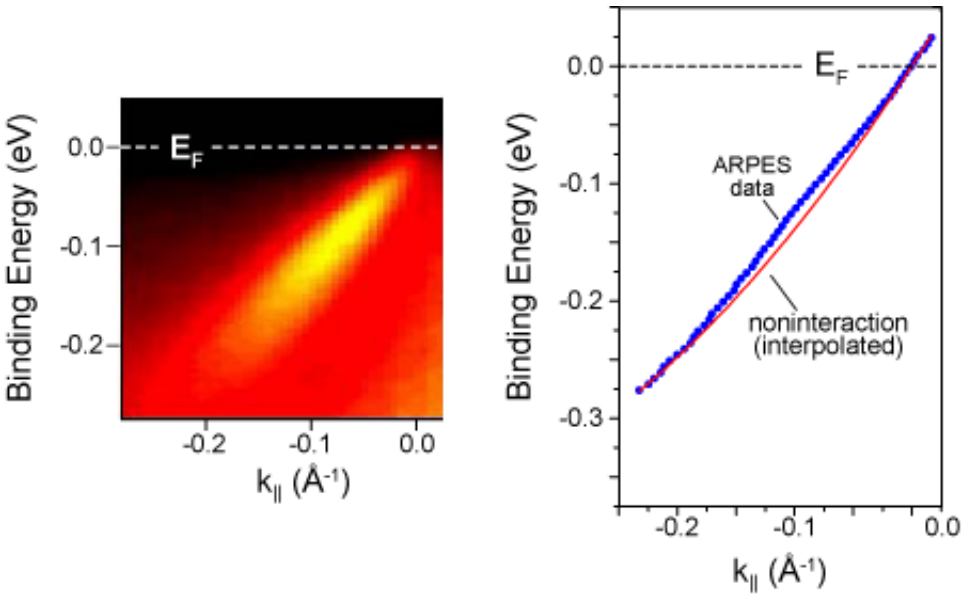


Figure 26: *ARPES data from the iron (110) surface state. Left: Raw data, showing the intense quasiparticle region. Right: The electron band dispersion (E vs. $k_{||}$) extracted from the data reveals a weak "kink" in the region between 0.1 and 0.2 eV below E_F. From [29].*

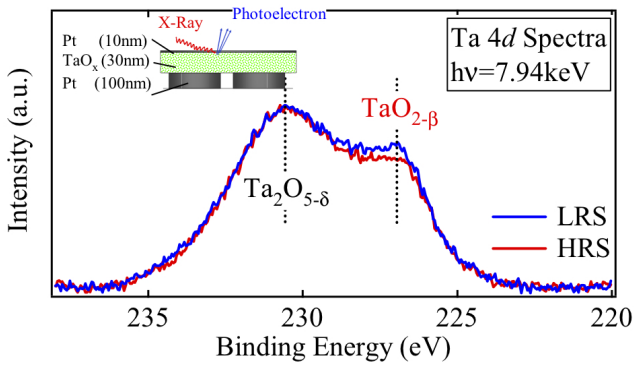


Figure 27: *Hard x-ray photoemission spectra from Ta_2O_5 in the high (HRS) and low resistive state (LRS). From [58].*

deviation from the parabolic behavior in the regime down to 200 meV below the Fermi level. This broader kink structure can be indeed related to the electron-magnon interaction [29]. From these data it is possible to extract the strength and extension of the electron magnon interaction.

4.5 High-Energy Photoemission (HAXPES)

So far we have discussed effects in valence band and core level photoelectron spectroscopy at excitation energies below 1000 eV. As we know from the inelastic mean free path curves under these conditions we will have $\lambda_{in} \simeq 1\mathrm{nm}$ at best, i.e. all of these experiments are more or less surface sensitive. In recent years there is a strong effort to extend photoelectron spectroscopy also to higher excitation energies up to 10 keV in order to overcome this limitation. The approach is coined HArd X-ray PhotoElectron Spectroscopy (HAXPES) and poses several experimental challenges [56]. First, the electron spectrometers must be modified to be able to measure photoelectrons with high kinetic energy and good energy resolution ($\Delta E < 100\,\mathrm{meV}$). Second, the photoexcitation cross section for most core levels drops by 2-3 orders of magnitude, when going from 1 keV to 10 keV photon energy. As a consequence, the resulting photoelectron intensity will be small and difficult to measure. This can be only partially compensated on the primary side, i.e. by increasing the photon flux. At present, HAXPES experiments are still demanding and very difficult to carry out with laboratory sources. With synchrotron radiation, however, HAXPES is quickly maturing into a powerful tool for materials characterization.

Laterally Integrating Spectroscopy – The major advantage of HAXPES is its larger information depth which permits the access to buried layers and interfaces. The example shown in Fig. 27 is taken from the field of resistive oxides. Usually, oxides are wide band gap insulators. As is discussed in Chapter E4, some of these materials may change their conductivity by several orders of magnitude, if a short current pulse above a certain threshold is applied to the material [57]. This current leads to the formation of conductive filaments or a local valency change in the oxide generating carriers for electrical transport. This is called the low resitive state (LRS). Interestingly, this process is reversible and the system may also be switched back into the high resistive state (HRS). This behaviour considered as a future memory principle and explains the strong interest in resistive oxides.
Ta_2O_5 is one of the promising materials that is currently thoroughly investigated with respect to resistive memory applications. Fig. 27 shows the comparison of photospectra taken from the Ta $4d$ core states with about 8 keV photon energy. In order to switch the conductivity of the Ta_2O_5 film a bottom and top electrode usually made from Pt is needed, through which the switching

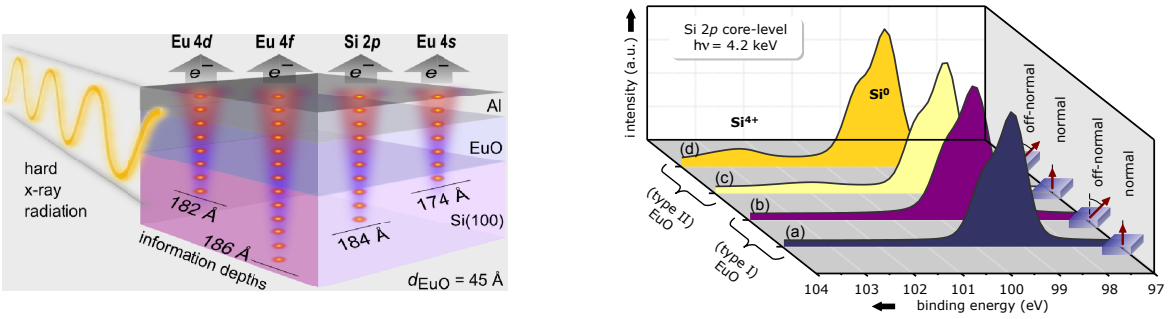


Figure 28: *HAXPES on the EuO/Si interface. (Left) Schematic representation of the information depth for the different Eu and Si core levels. (Right) Si 2p core level photoemission spectra for (type I) stoichiometric EuO and (type II) O-rich EuO, recorded at 4.2 keV photon energy in normal (0°) and off-normal (60°) electron emission geometry. From [61].*

current is passed through the insulator. This means, however, that the photoemission experiment has to probe the region below the Pt electrode, which requires a sufficiently high information depth. As can be seen, the experiment is indeed able to find a difference in the relative core level intensities underneath the 10 nm thick Pt-electrode, which can be related to a change of the oxidation state from Ta^{5+} to Ta^{4+} between the HRS and LRS state [58]. This demonstrates that HAXPES is able – at least in principle – to follow and map the valency changes taking place during the resistive switching process.

The second example relates to the field of spintronics. Magnetic tunneling barriers are considered as a means to enable an efficient spin injection into semiconductors [59, 60]. One of the materials for spin-filter barriers investigated in this context is the ferromagnetic semiconductor EuO. In order to obtain a well-defined system for spin injection, a chemically and structurally sharp interface between EuO and the semiconductor – preferably Silicon – must be established during the growth process. Of particular importance is the control of the oxygen partial pressure, as excess oxygen leads to a formation of interfacial silicon oxide.

The chemical quality of the EuO/Si interface can be addressed by HAXPES exploiting the kinetic energy dependence of the photoelectron inelastic mean free path (Fig. 28) [61]. The thickness of the EuO film ($d_{EuO} = 45$ Å) has been chosen such that for a given photon energy (4.2 keV) the photoelectrons from the Si 2p levels reaching the spectrometer originate mainly from the interfacial region between EuO and Si. The interface sensitivity can be even increased by changing the take-off angle of the electrons from normal emission to off-normal emission. As can be seen in Fig. 28 the growth of "oxygen-rich" EuO (type II) leads to a significant photoemission satellite in the Si 2p spectrum which stems from a Si^{4+}state. The spectral weight of this contribution increases for the off-normal emission geometry. This is a clear indication that the silicon oxide contribution is located at the EuO/Si interface. The growth of "stoichiometric" EuO takes place at a lower oxygen partial pressure. The respective Si 2p spectra prove the absence of an oxide component, i.e. the interface between EuO and Si is chemically sharp. The results for the Eu 4d, 4f, and 4s core level photoemission corroborate the findings.

Recent experiments have demonstrated that high-energy photoemission can also be used to carry out bulk-sensitive band mapping experiments [62–65].

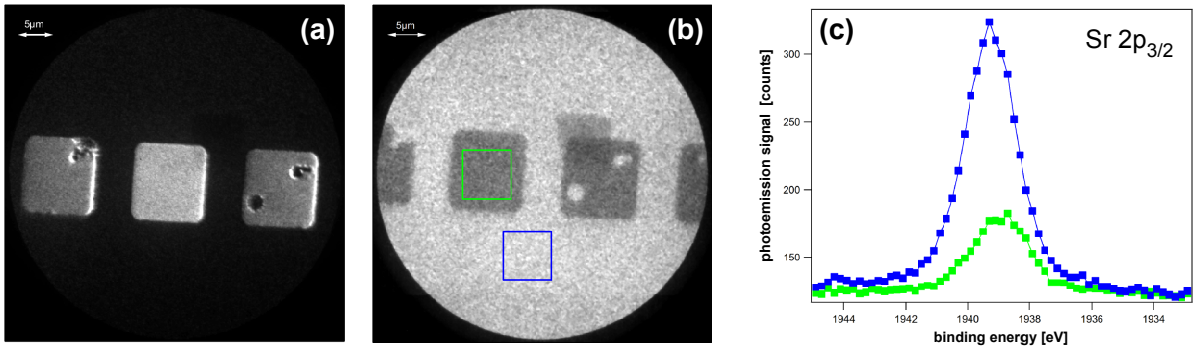


Figure 29: *Laterally resolved HAXPES on Au microstructures on $SrTiO_3$. (a) threshold photoemission, (b) Sr $2p_{3/2}$ photoelectron distribution at $h\nu = 6500$ eV, and (c) local photoelectron spectra from the regions of interest (blue, green).*

High-Energy Spectromicroscopy – The NanoESCA system offers a possibility to perform HAXPES experiments with lateral resolution. A first demonstration for a resisitive material is given in Fig. 29. The sample consists of 10 nm thick Au electrode micropatterns on a $SrTiO_3$ substrate, which show up as bright areas in threshold photoemission. At a photon energy of $h\nu = 6500$ eV, the kinetic energy of the $2p_{3/2}$ photoelectrons is high enough to allow them to penetrate the Au electrode. The respective photoelectron distribution recorded with the NanoESCA exhibits a weaker photoemission signal from the area of the electrodes (Fig. 29b) than from the uncovered part of the $SrTiO_3$ substrate. The lateral resolution in this experiment is about 500 nm. In order to obtain photoemission spectra, a sequence of images at different kinetic energies covering the photoemission peak has been acquired. This image stack is analyzed by integrating the intensity in each image for the two marked regions of interest (blue and green) and plotting it as a function of the kinetic energy. This analysis yields the local $2p_{3/2}$ photoemission spectra in Fig. 29c. The relative difference in the total intensities of the two spectra relates to the damping by the Au layer. One also notes, however, a slight shift of the peak position between the two spectra. This suggests that the chemical states of the covered and uncovered surface areas are slightly different. This experiment demonstrates that HAXPES can also performed with sub-micrometer lateral resolution.

4.6 Interfacial sensitivity

In Sect. 4.5 we have demonstrated two ways to vary the surface sensitivity in photoemission: changing the photon energy so as to move along curves of the type in Figure 5 and varying the take-off angle, as indicated e.g. in Figure 28. Both of these involve electron escape processes. One may also ask if there is a way to somehow tailor the photon wave field so as to vary surface sensitivity. Creating an x-ray standing wave is one method for doing this, and it has been found possible to selectively look at buried layers and interfaces [66], as well as element-resolved densities of states [67], in this way.

In Figure 30, we illustrate one approach for using soft x-ray (or in the future also hard x-ray) standing waves to carry out more precise depth-resolved photoemission from multilayer nanostructures [66]. This x-ray standing wave (XSW) approach combines a standing wave created by first-order Bragg reflection from a multilayer mirror of period d_{ML} with a sample in which one layer has a wedge profile, and can be termed the “swedge method”. If the standing

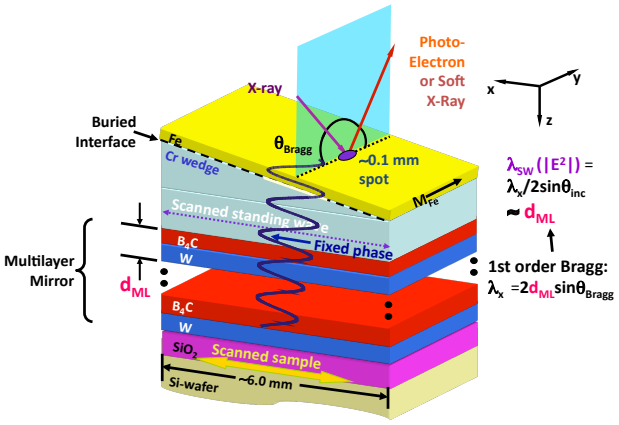


Figure 30: *Schematic illustration of the simultaneous use of an x-ray standing wave plus a wedge-profile overlayer sample to selectively study buried interfaces and layers – the "swedge" method. In the example here, a strong standing wave (SW) is created by first-order Bragg reflection from a multilayer made of repeated B_4C/W bilayers, and a Cr wedge underneath an Fe overlayer permits scanning the SW through the Fe/Cr interface by scanning the sample along the x direction. From ref. [66].*

wave is created by a typically well-focussed synchrotron radiation beam, then its dimensions will be much smaller than a typical sample, as indicated in the figure. Since the standing wave only exists in the region where the beam hits the sample surface, and its phase is locked tightly to the multilayer mirror, scanning the sample in the photon beam along the x direction effectively translates the standing wave through the sample. In the example shown, the standing wave would in particular scan through the Fe/Cr interface of interest, at some positions being more sensitive to the Fe side and at some more sensitive to the Cr side.

Some results obtained with this method for the Fe/Cr interface are summarized in Figs. 31 and 32. The analysis combined XPS intensity and MCDAD measurements (not shown here) from the 3p and 2p core levels of Fe and Cr, respectively. In Figure 31a is shown the variation of the Cr3p/Fe3p ratio as the sample is scanned in the way suggested above, for several angles of incidence near the Bragg angle. Oscillations in this ratio clearly reflect the passage of the standing wave node and belly through the interface. In Figure 31b we compile "rocking curves" in which the angle is varied around the Bragg angle for different positions x along the sample, or equivalently different Cr wedge thickness d_{Cr}. There are dramatic changes in the intensity ratio in this data also.

Self-consistently analyzing these data with x-ray optical calculations of standing-wave photoemission and only two variable parameters (the depth of onset of change in the Fe composition and the width of a linear gradient as the interface changes from pure Fe to pure Cr) yields the excellent fits shown to both types of data, and the parameters given at the left side of Figure 32a. The MCDAD data for both Fe 2p and Cr 2p core level photoemission have also been measured as the sample is scanned in the beam. The relative signs of the MCDAD signal for the Fe 2p and Cr 2p levels are found to be opposite [66]. this immediately implies that a small amount of Cr is oppositely magnetized compared to Fe, and this must be induced by the ferromagnetic Fe layer, since Cr is normally antiferromagnetic. Similar data have been obtained at the 3p levels of Cr and Fe. Further analyzing this data set with two parameters for Fe 2p and 3p MCD and two parameters for Cr 2p and 3p MCD yields the atom-specific magnetization profiles shown at right hand side of Figure 32a.

Thus, in the above described experiment the swedge method has permitted non-destructively determining the concentration profile through an interface, as well as the atom-specific magnetization contributions through it. The swedge approach has also been used successfully to determine layer-specific densities of states that can be linked changes in magnetoresistance as a function of nanolayer thicknesses [68]. Several other possible applications of it have also been

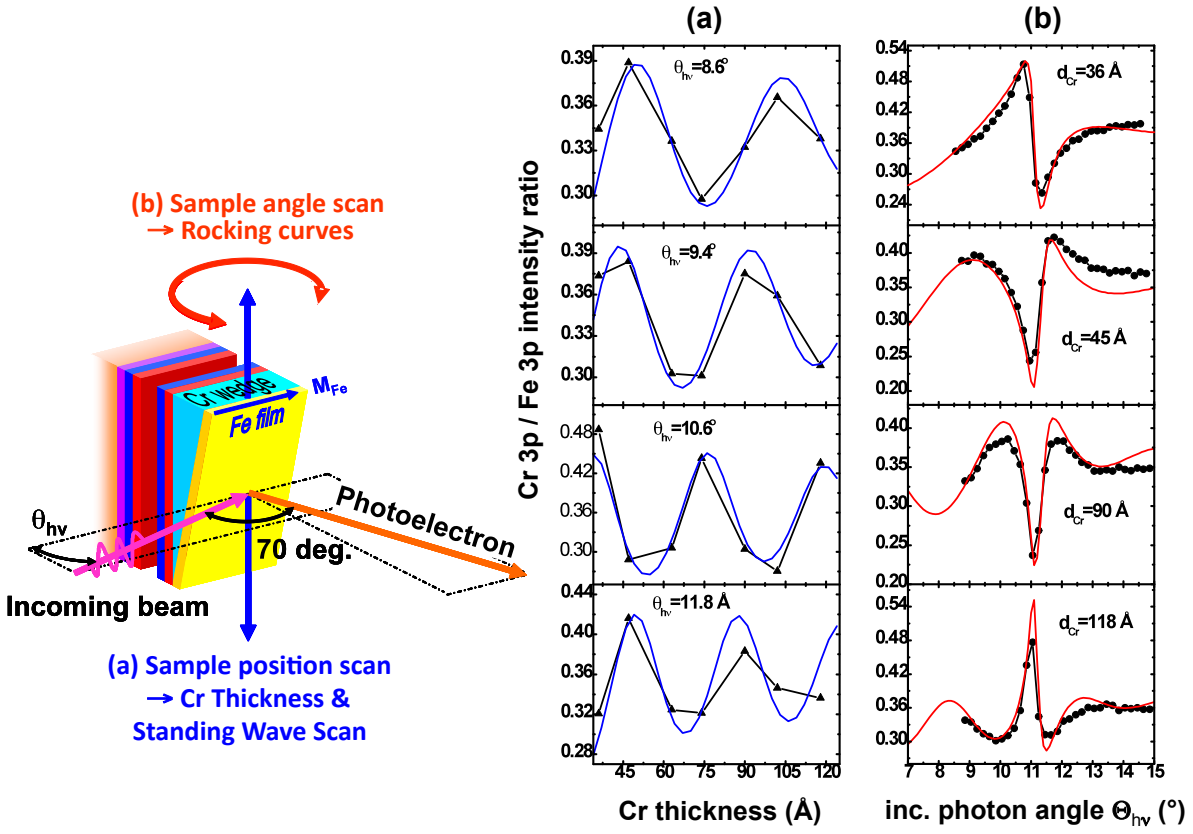


Figure 31: *Experimental and calculated Cr3p/Fe3p ratios for two types of standing wave scan for the sample shown in Figure 21: (a) Scanning the sample along x at fixed incidence angle, as indicated in Figure 30, and (b) scanning the sample polar angle with fixed x position (or Cr thickness). The best-fit theory curves are for the parameters shown at the left of Figure 24(a). From ref. [66].*

suggested [66, 69, 70], including going to hard x-ray excitation, for which reflectivities and thus standing wave strengths can be much higher.

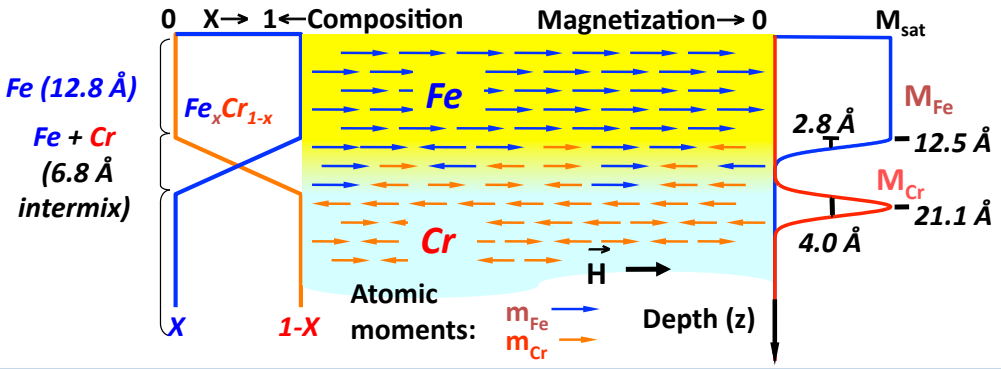


Figure 32: *The concentration and atom-specific magnetization profiles through the Fe/Cr interface, as derived from the XPS and MCDAD experiments. From ref. [66].*

5 Conclusions

In this contribution, we could only touch upon selected aspects of photoelectron spectroscopy and photoemission processes. It should have become clear that this spectroscopy with its many facets provides a powerful tool for an *en detail* electronic and chemical characterization of materials. Very important information can already be extracted by means of qualitative interpretation schemes. The full potential, however, can be unleashed by quantitative descriptions within sophisticated photoemission calculations. The successful expansion of photoemission techniques to hard x-ray excitation relaxes the constraint of surface sensitivity. HAXPES offers access to genuine bulk electronic structures and buried interfaces.

Acknowledgement

The author is indepted to the Jülich spectroscopy and microspectroscopy groups at the storage ring facilities DELTA (Dortmund), BESSY (Berlin) and ELETTRA (Trieste), particularly S. Cramm, S. Döring, V. Feyer, D. Gottlob, A. Herdt, I. Krug, F. Nickel, M. Patt, L. Plucinski, and C. Wiemann.
Sincere thanks are due to C. S. Fadley (ALS Berkeley), A. X. Gray (SLAC), A. M. Kaiser (SPECS GmbH), M. Merkel (Focus GmbH), and M. Escher (Focus GmbH) for ongoing collaboration and the permission to use material for this lecture.

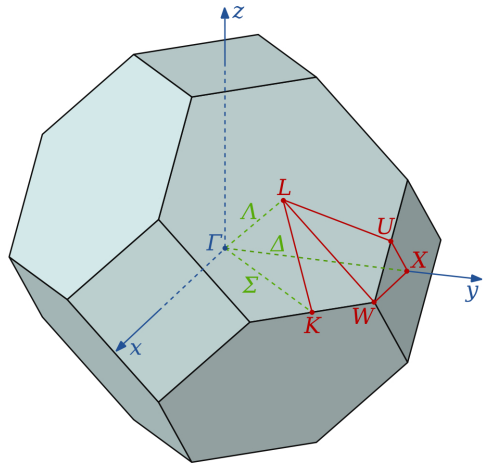


Figure 33: *Brillouin zone of an face-centered cubic crystal with high-symmetry directions (green) and high-symmetry points (red).*

Appendix I : Optical Selection Rules

The strength of the optical interband transitions in the electronic structure of a solid is determined by dipole selection rules. These dipole selection rules depend on the symmetry of the crystal and the experiment and have been derived from group-theoretical arguments. In the following, we will give an example for the dipole selection rules valid in a face-centered cubic crystal.

The symmetry of the electronic wavefunctions in the initial and final state of the photoemission process is described by so-called irreducible representations A_i of the symmetry group. The letter A usually refers to high-symmetry directions and points in the Brillouin zone (Fig. 33). Depending on the structure of the Hamiltonian (see, for example, eq. 2) one may distinguish *nonrelativistic* or *single-group representations* (spatial symmetry only), *relativistic* or *double-group representations* (including spin-orbit coupling) and *magnetic co-representations* (spin-orbit coupling and exchange interaction). The dipole operator of the photon field can also be expressed by an appropriate representation, taking into account the experimental geometry (direction of light incidence). For an fcc crystal, we have the following representations without and with spin-orbit coupling along the three high-symmetry directions

crystalline axis	**symmetry label**	**nonrelativistic**	**relativistic**
{001}	Δ	Δ_1, Δ_2,$\Delta_{2'}$,Δ_5	Δ_6^1, Δ_6^5, Δ_7^2, $\Delta_7^{2'}$, Δ_7^5
{011}	Σ	Σ_1, Σ_2, Σ_3, Σ_4	Σ_5^1, Σ_5^2, Σ_5^3, Σ_5^4
{111}	Λ	Λ_1, Λ_3	$\Lambda_{4,5}^3$, Λ_6^1, Λ_6^3

Table 1:

Electronic states with the highest symmetry have the lowest number, i.e. $\Delta_1, \Sigma_1, \Lambda_1$. In a photoemission process along the surface normal of a low-index surface, the electron is excited into a final state with the highest symmetry inside the crystal, which connects to a spherical wave (corresponding to a s wavefunction) outside the crystal propagating towards the detector. For this normal emission geometry, we have the following dipole selection rules between single-group representations [71]

In the presence of spin-orbit coupling, the interband transitions take place between double-group representations and lead to the excitation of spin-polarized photoemission due to the

	final state			
initial state	Δ_1	Δ_2	$\Delta_{2'}$	Δ_5
Δ_1	$\parallel$			$\perp$
Δ_2		$\parallel$		$\perp$
$\Delta_{2'}$			$\parallel$	$\perp$
Δ_5	$\perp$	$\perp$	$\perp$	$\perp$

	final state			
initial state	Σ_1	Σ_2	Σ_3	Σ_4
Σ_1	$\parallel$		X	Y
Σ_2		$\parallel$	Y	X
Σ_3	X	Y	$\parallel$	
Σ_4	Y	X		$\parallel$

	final state	
initial state	Λ_1	Λ_3
Λ_1	$\parallel$	$\perp$
Λ_3	$\perp$	$\parallel$,$\perp$

Table 2: *$\vec{A}$ parallel ($\parallel$) and perpendicular ($\perp$) to the surface normal, and parallel to the x-axis (X) or y-axis (Y), respectively.*

mechanism of optical spin-orientation. An example for the resulting relativistic selection rules with circularly polarized light are given in Table 3. For a full set of relativistic selection rules, the reader is referred to Ref. [72].

	final state				
initial state	Δ_6^1	Δ_7^2	$\Delta_7^{2'}$	Δ_6^5	Δ_7^5
Δ_6^1				$\|\Uparrow\rangle$	$\|\Downarrow\rangle$
Δ_7^2				$\|\Downarrow\rangle$	$\|\Uparrow\rangle$
$\Delta_7^{2'}$				$\|\Downarrow\rangle$	$\|\Uparrow\rangle$
Δ_6^5	$\|\Downarrow\rangle$	$\|\Uparrow\rangle$	$\|\Uparrow\rangle$		
Δ_7^5	$\|\Uparrow\rangle$	$\|\Downarrow\rangle$	$\|\Downarrow\rangle$		

Table 3: *Relativistic dipole selection rules for normal emission along the Δ line and excitation with right-hand circularly polarized light. The spin-polarization is oriented along the surface normal (z-axis) and can be positive ($|\Uparrow\rangle$) or negative ($|\Downarrow\rangle$). For excitation with left-hand circularly polarized light, $|\Uparrow\rangle$ and $|\Downarrow\rangle$ have to be interchanged.*

References

[1] S. Hüfner, *Photoelectron Spectroscopy* 3rd ed. (Springer, Berlin, 2003).

[2] *Solid-State Photoemission and Related Methods*, eds. W. Schattke, and M. A. van Hove (Wiley-VCH, Weinheim, 2003).

[3] F. Reinert and S. Hüfner, New J. Phys. **7**, 97 (2005).

[4] *Very High Resolution Photoelectron Spectroscopy*, ed. S. Hüfner (Springer, Berlin, 2007).

[5] F. de Groot and A. Kotani, *Core Level Spectroscopy of Solids* (CRC Press, Boca Raton, 2008).

[6] R. A. Millikan, Phys. Rev. **7**, 355 (1916).

[7] H. Hertz, Annal. Phys. **267**, 983 (1887).

[8] A. Einstein, Annal. Phys. **322**, 132 (1905).

[9] C. Nordling, E. Sokolowski, and K. Siegbahn, Phys. Rev. **105**, 1676 (1957).

[10] *Handbook on Synchrotron Radiation* Vol. 1A, B, edited by D. Eastman and Y. Farge (North-Holland Publishing, Amsterdam, 1983).

[11] M. Salmeron and R. Schlögl (2008). Lawrence Berkeley National Laboratory Paper LBNL-183E. Retrieved from: http://www.escholarship.org/uc/item/6n75v256.

[12] R. M. Martin, *Electronic Structure – Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[13] K. Horn, in *Handbook of Surface Science*; Vol. 2, edited by K. Horn and M. Scheffler (Elsevier, Amsterdam, 2000), p. 383.

[14] *Optical Orientation*, edited by F. Meier and B. P. Zakharchenya (North-Holland, Amsterdam, 1984).

[15] H. Ebert, J. Minar, and V. Popescu, in: *Band Ferromagnetism*; edited by K. Baberschke, M. Donath, and W. Nolting (Springer-Verlag, Berlin, 2001), p. 371.

[16] R. N. Zare, *Angular Momentum* (Wiley, New York, 1988).

[17] D. T. Pierce and F. Meier, Phys. Rev. B **13**, 5484 (1976).

[18] [ ] C. S. Fadley, in: *Synchrotron Radiation Research: Advances in Surface and Interface Science*, R. Z. Bachrach, Ed. (Plenum Press, New York, 1992).

[19] Multiple scattering program for calculating photoelectron diffraction available at: http://csic.sw.ehu.es/jga/software/edac/index.html, with the methodology behind it described in F.J. Garcia de Abajo, M.A. Van Hove, and C.S. Fadley, Phys. Rev. B **63**, 075404 (2001).

[20] S. Tanuma, C. J. Powell, and D. R. Penn, Surf. Interface Anal. **37**, 1 (2005); S. Tanuma, C. J. Powell, and D. R. Penn, "*Calculations of electron inelastic mean free paths. IX. Data for 41 elemental solids over the 50 eV to 30 keV range*", to be published.

[21] R. Feder, in: *Polarized Electrons in Surface Physics*, ed. by R. Feder (World Scientific, Singapore, 1985).

[22] G. Borstel, Appl. Phys. A **38**, 193 (1985).

[23] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003).

[24] A. Oelsner, O. Schmidt, M. Schicketanz, M.J. Klais, G. Schönhense, V. Mergel, O. Jagutzki, H. Schmidt-Böcking, Rev. Sci. Instrum. **72** (2001) 3968.

[25] N. Takahashi, F. Matsui, H. Matsuda, Y. Hamada, K. Nakanishi, H. Namba, and H. Daimon, J. Electron Spectr. Rel. Phen. **163**, 45 (2008).

[26] J. Kessler, *Polarized Electrons*, 2nd ed. (Springer-Verlag, Berlin, 1985).

[27] J. Kirschner, *Polarized Electrons at Surfaces*, Springer Tracts in Modern Physics Vol. 106 (Springer-Verlag, Berlin, 1985).

[28] L. Gregoratti, M. Marsi and M. Kiskinova, Synchr. Rad. News **12**, 40 (1999).

[29] M. Escher, N. Weber, M. Merkel, C. Ziethen, P. Bernhard, G. Schönhense, S. Schmidt, F. Förster, F. Reinert, B. Krömker, and D. Funnemann, J. Phys.: Condens. Matt. **17**, S1329 (2005).

[30] V. N. Strocov, R. Claessen, G. Nicolay, S. Hüfner, A. Kimura, A. Harasawa, S. Shin, A. Kakizaki, H. I. Starnberg, P. O. Nilsson, and P. Blaha, Phys. Rev. B **63**, 205108 (2001).

[31] available as pdf-file at `http://xdb.lbl.gov/`

[32] S. Doniach and M. Sunjic, J. Phys. C **3**, 285 (1970).

[33] E. Rotenberg, Advanced Light Source Berkeley, (priv. communication).

[34] S. V. Borisenko, M. S. Golden, S. Legner, T. Pichler, C. Dürr, M. Knupfer, J. Fink, G. Yang, S. Abell, and H. Berger, Phys. Rev. Lett. **84**, 4453 (2000).

[35] Y. J. Kim, C. Westphal, R. X. Ynzunza, Z. Wang, H. C. Galloway, M. Salmeron, M. A. Van Hove, C. S. Fadley, Surf. Sci. **416** (1998) 68.

[36] J. Osterwalder, A. Tamai, W. Auwarter, M.P. Allan, and T. Greber, Chimia **60** (2006) A795, and earlier references therein.

[37] A. Bailly, O. Renault, N. Barrett, T. Desrues, D. Mariolle, L. F. Zagonel, and M. Escher, J. Phys.: Condens. Matt. **21**, 314002 (2009).

[38] B. Kroemker, M. Escher, D. Funnemann, D. Hartung, H. Engelhard, and J. Kirschner, Rev. Sci. Instrum. **79**, 053702 (2008).

[39] E. Kisker, K. Schröder, W. Gudat, and M. Campagna, Phys. Rev. B **31**, 329 (1985).

[40] H. B. Rose, A. Fanelsa, T. Kinoshita, Ch. Roth, F. U. Hillebrecht, and E. Kisker, Phys. Rev. B **53**, 1630 (1996).

[41] E. Tamura and R. Feder, Europhys. Lett. **16**, 695 (1991).

[42] C. M. Schneider and J. Kirschner, Crit. Rev. Solid State Mater. Sci. **20**, 179 (1995).

[43] E. I. Rashba, Sov. Phys. Solid State **2**, 1109 (1960).

[44] S. LaShell, B. A. McDougall and E. Jensen, Phys. Rev. Lett. **77**, 3419 (1996).

[45] F. Reinert, J. Phys.: Condens. Matt. **15**, S693 (2003).

[46] W. Kuch and C. M. Schneider, Rep. Prog. Phys. **64**, 205 (2001).

[47] C. Bethke, E. Kisker, N. B. Weber, and F. U. Hillebrecht, Phys. Rev. B **71**, 024413 (2005).

[48] D. Venus, Phys. Rev. B **43**, 6144 (1993).

[49] F. Manghi, V. Bellini, J. Osterwalder, T. J. Kreutz, and P. Aebi, C. Arcangeli, Phys. Rev. B **59**, R10409 (1999).

[50] K. Ono, K. Shimada, Y. Saitoh, T. Sendohda, A. Kakizaki, T. Ishii, and K. Tanaka, J. Electron Spectr. Rel. Phen. **78**, 325 (1996).

[51] Y. Sakisaka, T. Komeda, M. Onchi, H. Kato, S. Masuda, and K. Yagi, Phys. Rev. B **36**, 6383 (1987).

[52] F. Manghi, V. Bellini, and C. Arcangeli, Phys. Rev. B **56**, 7149 (1997).

[53] R. Clauberg, W. Gudat, E. Kisker, E. Kuhlmann, and G. M. Rothberg, Phys. Rev. Lett. **47**, 1314 (1981).

[54] G. Grimvall, *The Electron-Phonon Interaction in Metals*, Selected Topics in Solid State Physics, edited by E. Wohlfarth (North-Holland, New York, 1981).

[55] S.-J. Tang, J. Shi, B. Wu, P. T. Sprunger, W. L. Yang, V. Brouet, X. J. Zhou, Z. Hussain, Z.-X. Shen, Z. Zhang, and E. W. Plummer, phys. stat. sol. (b) **241**, 2345 (2004).

[56] K. Kobayashi, Nucl. Instrum. Methods A **601**, 32 (2009).

[57] R. Waser and M. Aono, Nat. Mater. **6**, 833 (2007).

[58] H. Kumigashira (priv. communication).

[59] R. Meservey and P. M. Tedrow, Phys. Rep. **238**, 173 (1994).

[60] G.-X. Miao, M. Münzenberg, and J. S. Moodera, Rep. Prog. Phys. **74**, 036501 (2011).

[61] C. Caspers, M. Müller, A. X. Gray, A. M. Kaiser, A. Gloskovskii, C. S. Fadley, W. Drube, and C. M. Schneider, Phys. Status Solidi RRL **1**, 1 (2011).

[62] L. Plucinski, J. Minar, B. C. Sell, J. Braun, H. Ebert, C. M. Schneider, and C. S. Fadley, Phys. Rev. B **78**, 035108 (2008).

[63] F. Venturini, J. Minar, J. Braun, H. Ebert, and N. B. Brookes, Phys. Rev. B **77**, 045126 (2008).

[64] C. S. Fadley, J. Electron Spectr. Rel. Phen. **178–179**, 2 (2010).

[65] A. X. Gray, C. Papp, S. Ueda, B. Balke, Y. Yamashita, L. Plucinski, J. Minar, J. Braun, E. R. Ylvisaker, C. M. Schneider, W. E. Pickett, H. Ebert, K. Kobayashi, and C. S. Fadley, Nat. Mater. **10**, 759 (2011).

[66] S.-H. Yang, B. S. Mun, N. Mannella, S.-K. Kim, J. B. Kortright, J. Underwood, F. Salmassi, E. Arenholz, A. Young, Z. Hussain, M. A. Van Hove, and C. S. Fadley, J. Phys. Cond. Matt. **14**, L406 (2002).

[67] J. Woicik, Nucl. Instr. Meth. A **547**, 227 (2005).

[68] S.-H. Yang, B. S. Mun, N. Mannella, A. Nambu, B. C. Sell, S. B. Ritchey, F. Salmassi, S. S. P. Parkin, C. S. Fadley, J. Phys.: Condens. Matter **18**, L259 (2006).

[69] C. S. Fadley, S.-H. Yang, B. S. Mun, J. Garcia de Abajo, in: *Solid-State Photoemission and Related Methods: Theory and Experiment*, W. Schattke and M.A. Van Hove (Eds.), (Wiley-VCH Verlag GmbH, Berlin, 2003).

[70] S.-H. Yang, B.S. Mun, and C.S. Fadley, Synchrotron Radiation News **24**, 17 (2004).

[71] W. Eberhardt and F. J. Himpsel, Phys. Rev. B **21**, 5572 (1980).

[72] M. Wöhlecke and G. Borstel, Phys. Rev. B **24**, 2857 (1981); M. Wöhlecke and G. Borstel, Phys. Rev. B **23**, 980 (1981); G. Borstel and M. Wöhlecke, Phys. Rev. B **24**, 2321 (1981).

F 4 X-ray Absorption Spectroscopy

L. Baumgarten
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

1 Introduction

In x-ray absorption spectroscopy (XAS) the sample is irradiated with monochromatic x-rays of energy hν. One scans the radiation energy and observes the absorption. When the energy is sufficient to excite an electron from a core level to above the Fermi level a jump in the absorption cross section is observed. Discrete structures can be observed just above the threshold. They are divided into the x-ray absorption near edge structure (XANES) and the extended x-ray absorption fine structure (EXAFS). In an atom one expects a few sharp (discrete) lines just above the threshold, followed by a continuum. XANES is a probe of the local environment of an atom or of the unoccupied density of electronic states (in accordance with selection rules).

Traditionally XAS spectra are measured in transmission mode. The spectrum is given by the ratio of the intensity before and after the absorber. Since a core hole created by the absorption of a x-ray photon will decay either by x-ray emission (emission of a photon of lower energy) or by emission of an Auger electron, both decay channels can as well used as a measure of the absorption strength. For low x-ray photon energies the preferred method of detection is that of the photo yield. If the absorption length is much larger as the escape depth of electrons or fluorescent radiation, this signal is proportional to the fraction of the incident radiation absorbed in the surface region.

X-ray absorption spectroscopy is element specific and can likewise be applied to gases, liquids, and solids. Almost any material like catalysts, minerals, and biological tissue (in its natural aqueous environment) can be investigated. It can be used to study surfaces, interfaces, buried layers and impurities at low concentrations. Special techniques have been developed to study magnetic layers and magnetic nanostructures, which are now widely used in magnetic storage devices and magnetic sensors.

X-ray absorption needs intense tunable X-rays. Therefore, today, the synchrotron is exclusively used as a source [1]. Other like XPS and Auger Spectroscopy it does not depend on ultrahigh vacuum conditions. In contrast to XPS which due to the limited escape depth of the photoelectrons is restricted to near surface areas it is feasible to also investigate subsurface structures like buried interfaces in semiconductors devices or GMR-sensors. A variety of books have appeared on the field of which only a few can be mentioned [3-7]. Varies aspects of this contribution have also been dealt with in earlier Spring Schools of this series [8-14]. The historical development of the field is described in ref. [15].

In the next chapter I will recall some basic X-ray features, then in chapter 3 I will discuss the near edge structures (XANES). The XANES, also called NEXAFS, probes the projected electronic density of empty states (local partial density of States LPDOS) and the valence of absorber atoms. In chapter 4 the EXAFS, which probe the interatomic distances and the coordination number will be briefly described. Chapters 5 and 6 are devoted to X-ray magnetic circular dichroism (XMCD) and x-ray magnetic linear dichroism (XMLD). These address special methods that have been developed to study magnetic films and nanoparticles. They are the prerequisite for the photoelectron emission microscopy (PEEM), which allows the study of magnetic domains and the dynamics of magnetic particles on the nm scale [16].

2 Some X-ray Basics

Traditionally an X-ray absorption spectrum is obtained from the ratio of the beam intensities in front and behind the absorber while the photon energy is tuned. Since X-rays are electromagnetic waves their absorption in a homogeneous material is described by the Lambert-Beer's law:

$$I = I_0 e^{-\mu x}. \quad (1)$$

Here I_0 is the incident X-ray intensity and I is the remaining intensity after the beam has propagated the distance x inside the sample. The linear absorption coefficient $\mu = \mu(n\omega)$ depends on the energy of the x-ray photons and is related to the absorption cross section $\sigma(n\omega)$ [$cm^2/atom$] as

$$\mu = \frac{\rho N_A}{A}\sigma \quad (2)$$

where ρ is the density of the target material, N_A is Avogadro's number and A is the atomic weight of the target material. To get a quantity that is independent on the target density the mass attenuation coefficient $\mu_m = \mu/\rho$ is often given. For mixed targets the mass attenuation coefficient can be calculated by summation of the contribution of the individual constituents.

$$\frac{\mu}{\rho} = \sum_i \left(\frac{\mu}{\rho}\right)_i c_i \quad (3)$$

where c_i is the weight concentration of the respective absorber atoms.

Four basic effects can contribute to the X-ray attenuation. These are the photoelectric absorption, the elastic (Rayleigh, Thompson) and the inelastic (Compton) scattering, and at energies above 1.022 MeV the electron-positron pair production. In the spectral range of interest here the photoelectric excitation is the dominant process, and we will consider only this contribution. Fig.1 shows the absorption coefficient at the L and K edges of a Cu-foil and in more detail the absorption fine structure in the vicinity of the K-edge.

In most of the spectral range plotted the attenuation coefficient decreases with increasing photon energy, approximately $\propto 1/(n\omega)^3$, except for certain points where it raises drastically. It is the vicinity of such "absorption edges" that we will study in some detail below. Except for the edges, the absorption coefficient is a smooth function of energy, with a value that depends on the sample density, the atomic number Z, atomic mass A, and the x-ray energy E roughly as

$$\mu \approx \frac{\rho Z^4}{AE^3} \quad (4)$$

For practical purposes it is often fit to so called victoreen functions [17] of the form

$$\mu(\lambda) = a \cdot \lambda^3 - b \cdot \lambda^4 \quad (5)$$

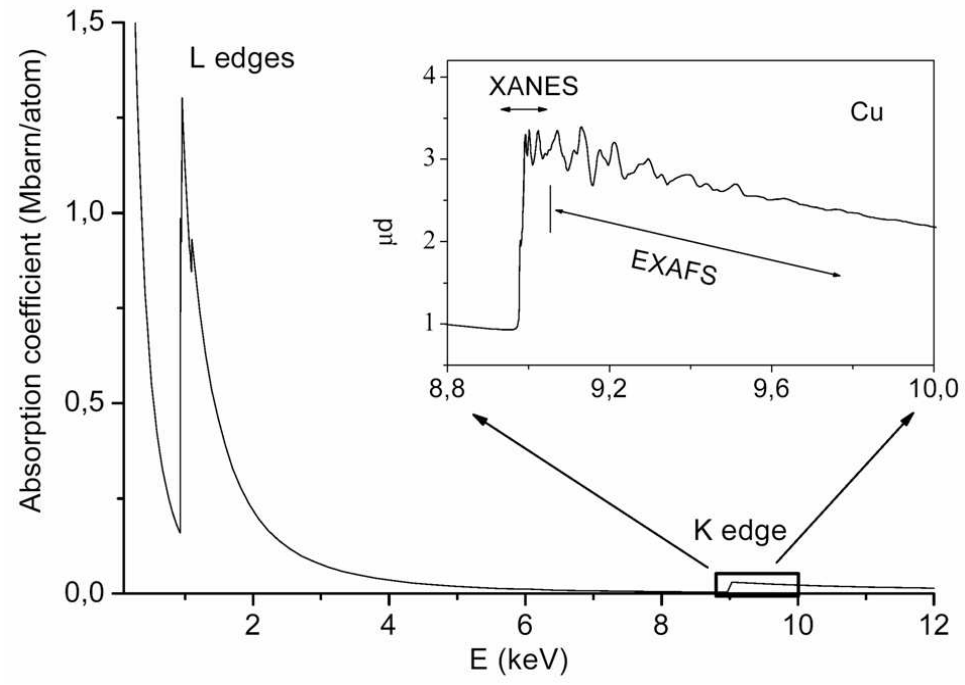


Fig. 1: *X-ray absorption coefficient of copper in the region of the* L *and* K *edges. The inset shows an expanded view of the* K *-edge region with the separation into a XANES and an EXAFS range [30].*

At the edges the photon energy reaches a value that is sufficient to excite an electron from a strongly bound core level to some high energy unoccupied state. With increasing energy it will even be ejected beyond the ionization limit into the continuum of states. The edges are denoted by letters K, L, M, . . . beginning this sequence with the edge of the highest energy. The K -edge is assigned to the excitation of a 1s-core level electron, the L_1 -edge to excitation of a 2s-electron, the L_2 and L_3 -edges to spin-orbit split $2p_{1/2}$ and $2p_{3/2}$ levels and so on for $M_1, \ldots, M_5, \ldots$.

The inset in fig. 1 shows the absorption coefficient in the vicinity of the K -edge of a Cu-foil in more detail. Such spectra are usually decomposed into three parts: The first part is the so called ”pre-edge” region which contains the background due to lower energy absorption transitions and sometimes also weak but specific absorption features to be discussed below. The second part is the XANES (X-ray Absorption Near Edge Structure) or NEXAFS (Near-Edge X-ray Absorption Fine Structure) - which basically means the same and includes the actual edge and the area to about 30eV beyond the edge. Because of the dipole selection rules XANES (NEXAFS) probes the projected electronic density of empty states (local partial density of States LPDOS) and the valence of the absorber atoms.

The third region is the EXAFS (Extended X-ray Absorption Fine Structure) range, which may extend to about 1000eV beyond the edge and essentially contains structural information about the immediate neighborhood of the absorber atom. EXAFS probes the coordination

number and the interatomic distance.
In the literature the basics of the excitation processes are often visualized in terms of a simple single particle picture (independent electron approach). Although this is convenient for discussion one has to be careful with such a model. For a quantitative interpretation and an understanding of finer details a consideration of electron correlation effects is indispensable.
The simple single particle picture cannot even properly explain the splitting of the 2p levels (L_2 and L_3 edges). The reason lies in the fact that the spin-orbit splitting of the 2p-levels is a "final state effect" rather than an "initial state effect"; that is to say, that in the initial state the 2p-electrons are in a $2p^6$ closed sub-shell configuration. In this closed shell configuration the angular momentum is completely canceled and, thus, there is no spin-orbit interaction. To understand the level splitting we have to argue in a more proper configuration picture [19] as follows: X-ray transitions - like other optical transitions are usually electric dipole transitions [1]. We therefore have to consider the following selection rules for excitations to be allowed:

$\Delta l = \pm 1$	$\Delta L = 0, \pm 1$
$\Delta m_l = 0, \pm 1$	$\Delta S = 0$
$\Delta m_s = 0$	$\Delta J = 0, \pm 1,$
$\Delta j = 0, \pm 1$	not allowed $0 \longleftrightarrow 0$

For the 2p levels p → s and p → d transitions are allowed. Due to Fermis Golden Rule the transition probability per unit time from an initial state |i> to a final state |**f**> is

$$P_{if} = \frac{2\pi}{\hbar} |\langle f | \boldsymbol{\epsilon} \cdot \mathbf{r} | i \rangle|^2 \rho_f(E) \tag{6}$$

where $\boldsymbol{\epsilon}$ is the polarization vector of the electromagnetic wave and $\boldsymbol{\epsilon} \cdot \mathbf{r}$ is the interaction operator in the dipole approximation. $\rho_f(E)$ is the energy density of final states. This equation is derived from time dependent perturbation theory for an atom that is loosely coupled to a classical radiation field [20]. The X-ray field is treated as a classical plain wave.
In a metal the edge transitions occur to final states just above the Fermi level [2]. For a transition metal, like Pt, with an incompletely filled d-electronic shell we therefore expect a dominance of the d-final states over the s-final states. Because of this dominance of d states we will ignore here the p → s transitions. Thus, in a configuration picture the initial state can be described by a $2p^6 5d^9$ configuration. Alternatively, because of the cancellation of the angular momenta in a completely filled shell, this initial state can also be described as a d^1 hole configuration of the d states. The final state $2p^5 5d^{10}$ has a closed d-shell and can be described as a p^1 hole configuration of the 2p states. A $L_{2,3}$ -edge transition is then a d → p hole transition from a

[1] Quadrupole transitions are occasionally observed but weak in intensity.

[2] The Pauli-exclusion principle does not allow electrons to be excited into occupied states.

[3] Here again we have a single particle picture. This time with a d hole ground state and a p hole excited state. In
[2] The Pauli-exclusion principle does not allow electrons to be excited into occupied states.

$5d^1$ to a $2p^1$ hole configuration [3]. In this final state we have an incompletely filled 2p shell, which naturally will show spin-orbit interaction. The spin- orbit splitting of the p shell is therefore a "final state effect". In a Russell-Saunders coupling scheme the L_3 - and L_2 -edges then correspond to $^2D \longrightarrow {}^2P_{3/2}$ and $^2D \longrightarrow {}^2P_{1/2}$ transitions, respectively. Note, that in this scheme the $\mathbf{J} = 3/2$ state has the lower energy as required by the Hund's rules [19]. Despite this described shortcoming of the single particle picture it is widely used in the literature to display the X-ray transitions, and we will also do so in the following.

The energetic position of the individual absorption edges increases with the atomic number approximately according to Moseleys law [21]:

$$\nu = C(Z - \sigma)^2 \tag{7}$$

where σ is a screening constant ($\sigma = 1.13$ for the K -shell and $\sigma \approx 7.9$ for the L - shell). That is, the higher the number of nuclear charges the stronger bound are the core electrons and the more energy will be needed to excite an electron to an empty valence state. This and the fact that the energies are so well defined makes X-ray spectroscopy so element specific.[4]

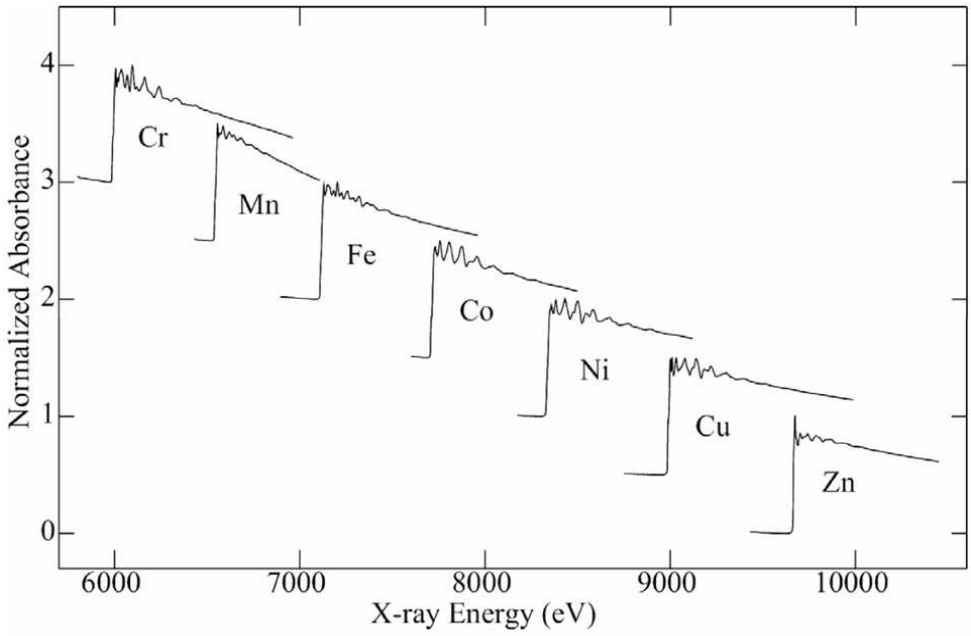


Fig. 2: *X-ray* K *-edge absorption of some* 3d *transition-metal foils. This figure clearly proofs the elemental selectivity of X-ray absorption. Above the edges the spectra clearly show some of the fine structure which is the subject of this contribution [23].*

[3] Here again we have a single particle picture. This time with a d hole ground state and a p hole excited state. In this picture the energy scale is turned upside down.

[4] Using the relationship (7) Dauvillier and Urbain in 1922 were able to identify element 72 (Hafnium) by its L emission lines [22]

Fig. 2 shows K-edge absorption spectra of the 3d metals Cr to Zn. It clearly demonstrates the shift of the absorption edge with atomic number. The figure also displays some of the fine structure that we want to discuss.

Once created, a core hole will rapidly decay either by X-ray fluorescence or by an Auger transition. Thereby the core hole is refilled by an electron from a higher level. In the first process a characteristic X-ray Photon is simultaneously emitted. As an optical transition it also obeys dipole selection rules. The photon energy matches the energy difference of the two levels and is characteristic to the atom. The emission lines are denoted by the signature of the initial core hole (K,L,M,. . .) with a numbered Greek index identifying the shell and the level from which the transition electron originates: K_{α_1,α_2} , K_{β_1}, ...

As an example Fig. 3 shows levels and transition lines for Cu (Z= 29). The position of the corresponding absorption edges which always are higher in energy are given as well [24, 25]. In the Auger process the excess energy is taken away by another electron [5]. Note that the

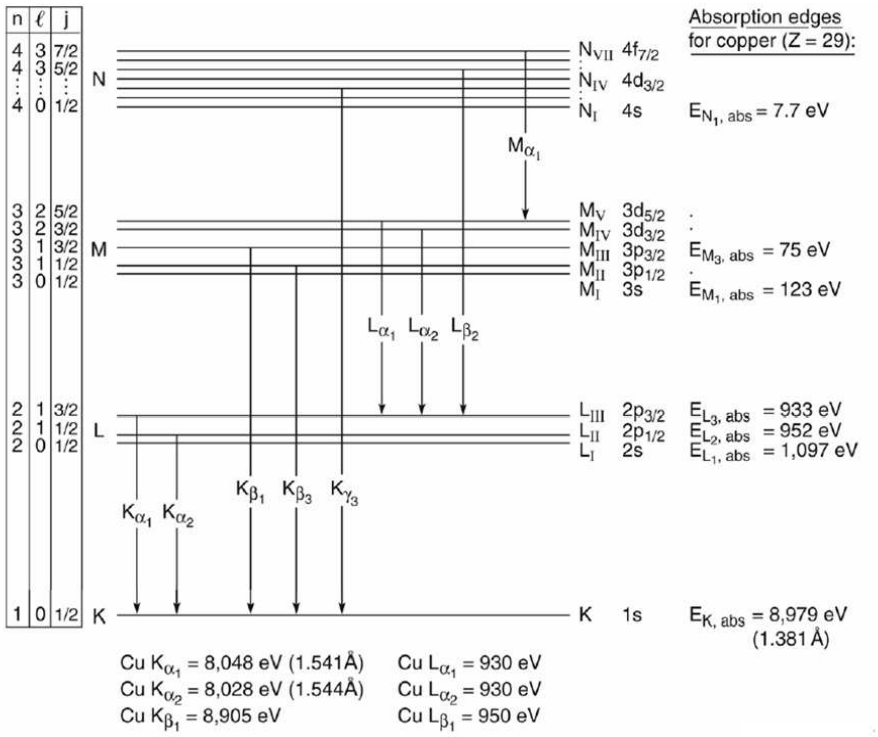


Fig. 3: *Correspondence of X-ray emission lines and participating atomic levels for the case of Cu (Z=29). The position of the absorption edges are also given [24, 25].*

Auger transition is not a two step process, but a single quantum mechanical transition. Auger transitions are characteristic to the atom as well. They are not restricted by dipole selection rules. Auger transitions are denoted by the three electronic levels involved, e.g. $K\,L_{2,3}\,M_{2,3}$, where the first letter labels the initial core hole. They are further classified as Coster-Kronig

[5] A review of the first 70 years of Auger spectroscopy has recently been given by W. Mehlhorn [26].

transitions, if one of the final-state holes has the same principle quantum number as the initial-state hole, e.g., $L_1 L_{2,3}M$, and super Coster-Kronig transitions, if both final-state holes have the same principle quantum number as the initial state hole, e.g. $L_1 L_{2,3} L_{2,3}$. Their energetics does not depend on the excitation process. This provides us with an easy means to distinguish Auger electrons from directly emitted photoelectrons. By changing the excitation wavelength the kinetic energy of direct photoelectrons will be shifted whereas the kinetic energies of Auger electrons are fixed. Such experiments can even be done at laboratory XPS - machines because usually two characteristic X-ray line sources are available, namely Al K_α at 1486.7 eV and Mg K_α at 1253.6 eV. The kinetic energy of the Auger electron can be estimated as $E_{kin} = E_{core} - E_1 - E_2$. Here E_{core}, E_1 , E_2 are the binding energies of the initial core hole and the other two electrons taking part in the process, respectively. Corrections will be necessary to this equation, however, because of correlation effects [27,28]. X-ray fluorescence and Auger transitions are competitive processes. Their relative abundance depends on the atomic number as shown in Fig.4. where the relative yield is plotted as a function of atomic number. For K -shell excitations the Auger transitions are dominant at $Z \leq 31$ (Ga) whereas the fluorescent emission is stronger at the heavier elements. For L-shell excitation the Auger transitions are dominant for all stable elements. Both processes, X-ray fluorescent emission and Auger decay themselves form the basis for widely used spectroscopies. Of course, Auger spectroscopy requires the better vacuum conditions. Due to the smaller scattering length of the ejected electrons it is also more surface

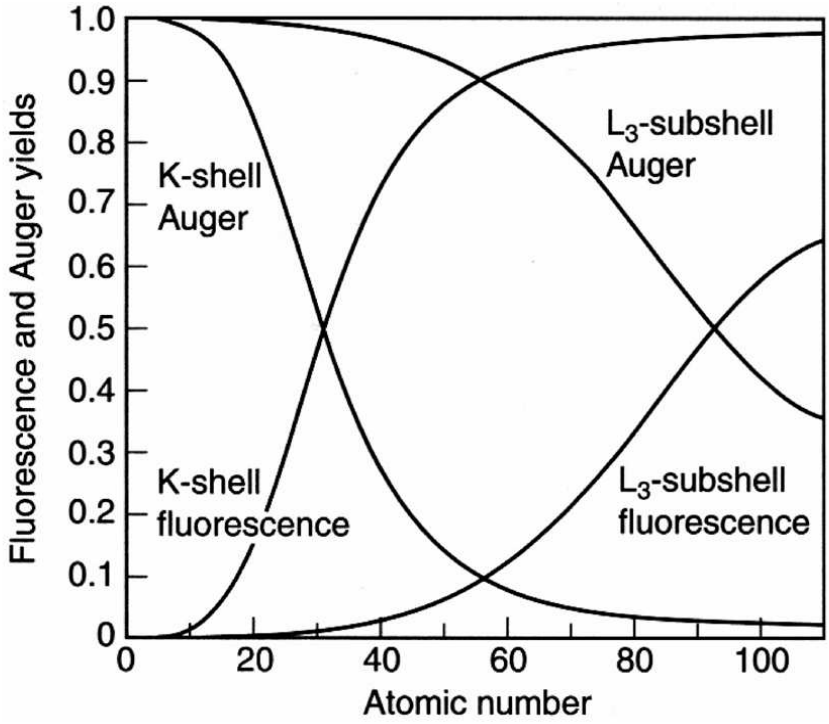


Fig. 4: *Comparison of X-ray fluorescence- and Auger yield as a function of atomic number for the* K *and* L_3 *absorption edges [25].*

sensitive than both, X-ray fluorescence spectroscopy and X-ray absorption spectroscopy.

For X-ray absorption the secondary processes imply the possibility of new detection schemes. An energetic Auger electron will often produce an avalanche of secondary electrons. A fraction of these will have enough energy to overcome the work function and leave the sample. Therefore, in the most easy experimental detection scheme the sample is simply connected to a pico-amperemeter and the drain current measured. This total-electron-yield (TEY) detection is a popular technique in solid state physics and surface science. The short scattering length of excited electrons limits the probing depth of this technique to about 10 nm. It is larger, though, than in XPS- and Auger measurements because here all the electrons are simultaneously collected, also the multiply scattered secondary ones. Alternatively, the emitted electrons and also the fluorescent photons may be measured by a suitable electron or photon detector in front of the sample. By choosing grazing incidence in reflection for the initial photon beam both detection techniques can be made surface sensitive. The distance that an electron can travel through the solid, the escape depth, varies with energy. Therefore, surface sensitivity can also be achieved by choosing only Auger electrons with a small escape depth for detection (partial-yield detection).

Fig. 5 shows a scheme of the experimental arrangement for a typical absorption measurement. In transmission mode a thin sample is placed into the monochromatized synchrotron beam. The beam intensity is monitored in front of and behind the sample by ionization chambers. It is common practice to measure the transmission of a reference sample in a third ionization chamber to eliminate fluctuations and to reliably detect small shifts of the samples absorption edge with respect to the reference. Fluorescence detection is preferentially used for very diluted absorbers because then the signal in transmission mode is the difference of two large intensities resulting

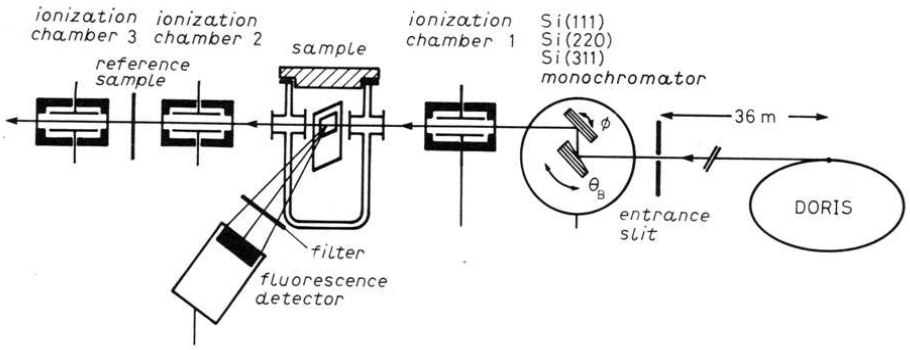


Fig. 5: *Arrangement of a typical X-ray absorption experiment[8,29].*

in large background noise. Also the reabsorption of the fluorescence photons is minimized in this mode. Energy dispersive Si(Li)- detectors, szintillation counters or gas filled proportional counters may be used to discriminate the sample fluorescence from the background signal. Photodiodes, channeltrons and channel plates are also used for detection.

3 Near Edge and Pre-edge Structures

The XANES part starts at the first strong dipole allowed transition (s → p transition in case of a K -absorption edge) below the ionization limit. This is nicely demonstrated in the gas phase spectrum at the K absorption edge of Ne in Fig. 7.

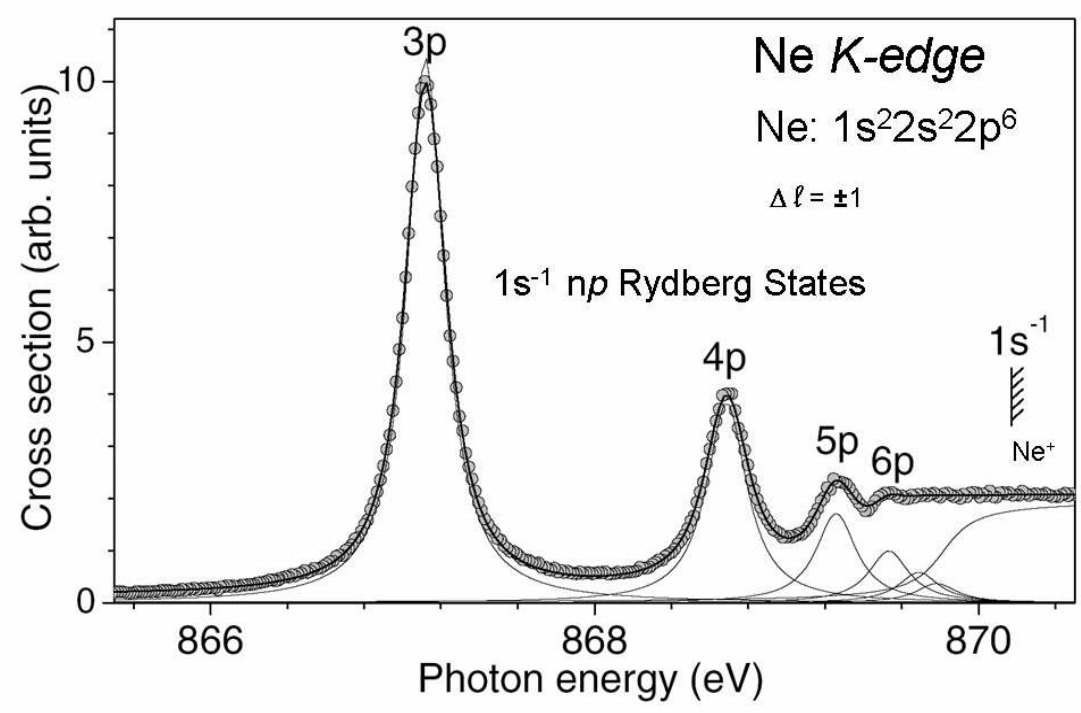


Fig. 7: *total cross section for photoionization of Ne in CO_2 at* K *edge absorption. The curves are fits to the data with a Gaussian with FWHM fixed to 66 meV [31].*

The Ne atom has a $1s^2 2s^2 2p^6$ configuration in its neutral ground state. Only transitions into unoccupied states are allowed. Therefore the first available unoccupied state in accordance with the $\Delta l = \pm 1$ selection rule is of 3p character, the next 4p, and so on. The ionization limit is also given. Note, that these levels correspond to $1s^{-1}np$ configurations where the negative exponent denotes the core hole. They do not give the positions of the atomic p levels but will be shifted due to correlation effects (screening).
The situation gets more complicated at the L-edges. Fig.8 shows the excitation at the L-edge of atomic Ar. Starting at the 2p levels the $\Delta l = \pm 1$ selection rule allows transitions to s and d levels. The first allowed transitions are to 4s and 3d. Due to the spin-orbit splitting of the 2p levels each peak appears twice. This is depicted in the level scheme on the right (but see also the discussion on the single particle picture above). The intensity ratio of corresponding levels is $\approx$ 2 : 1 because the electron occupation of the 2p levels is 4:2 in favor of the $2p_{3/2}$ levels. Spectra like this are often used for monochromator calibration.

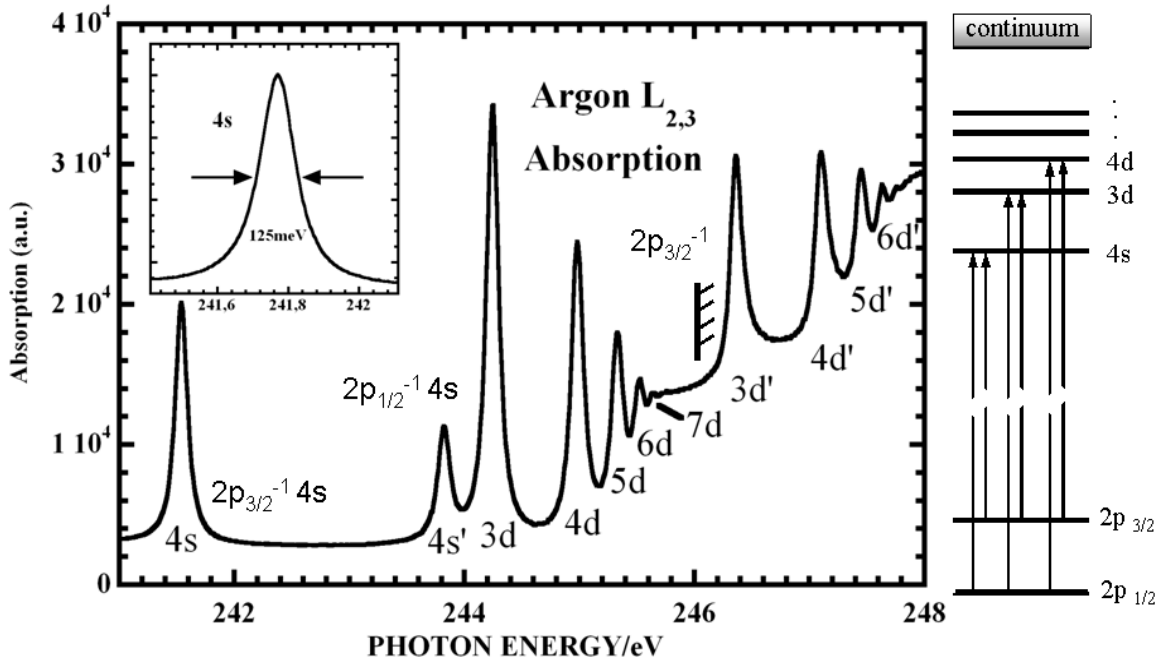


Fig. 8: *X-ray absorption spectrum of argon gas in the vicinity of the* $L_{2,3}$ *absorption edges and interpretation of the transitions in terms of the single particle picture. Such spectra are used to calibrate monochromators. [Courtesy of Stephan Cramm [1]]*

One might think that beyond the ionization limit of the 1s shell the X-ray absorption spectra of atoms will be smooth and flat as in Fig. 1. One rather observes multielectron excitations. These occur due to correlation effects. A typical example is a resonant double excitation process. The core electron is ejected by the photon and concomitantly a valence electron is excited to an empty valence orbital or ejected from the atom. In the first case we call it a "shake-up" process and in the second case a "shake-off" process. The shake- up and shake- off processes occur concurrently to the direct photoemission process. Of course, the total energy has to be conserved. Photoelectron waves passing through these two different pathways interfere with each other which results in an asymmetric line profile of the resonance in the photoabsorption cross section. They are called Fano resonances after U. Fano who predicted this lineshape [32]. One might see here a certain analogy to a double slit experiment. The electron has two optional pathways to go (this time on the energy scale) which leads to an interference pattern.

Fig.9 shows as an example 1s2s and 1s2p resonant double excitation spectra of Ne. The fit to the first peak shows a typical Fano profile [30, 31].

Fig.10 shows a high resolution K -shell absorption spectrum of N_2 in the gas phase. Above the ionization limit ($\approx$ 410 eV) besides the multi-electron excitations this shows a broad "shape" resonance. Such resonances are absent for isolated atoms. They are attributed to quasi-bound states in the continuum in which photoelectrons with high angular momentum about the center of the molecule are trapped due to multiple reflections in a centrifugal barrier potential

$l(l+1)/2r^2$ which reaches into the continuum states. Since the energy and the width of the resonances depend on the shape of the potential, they are called shape resonances [37]. Mul-

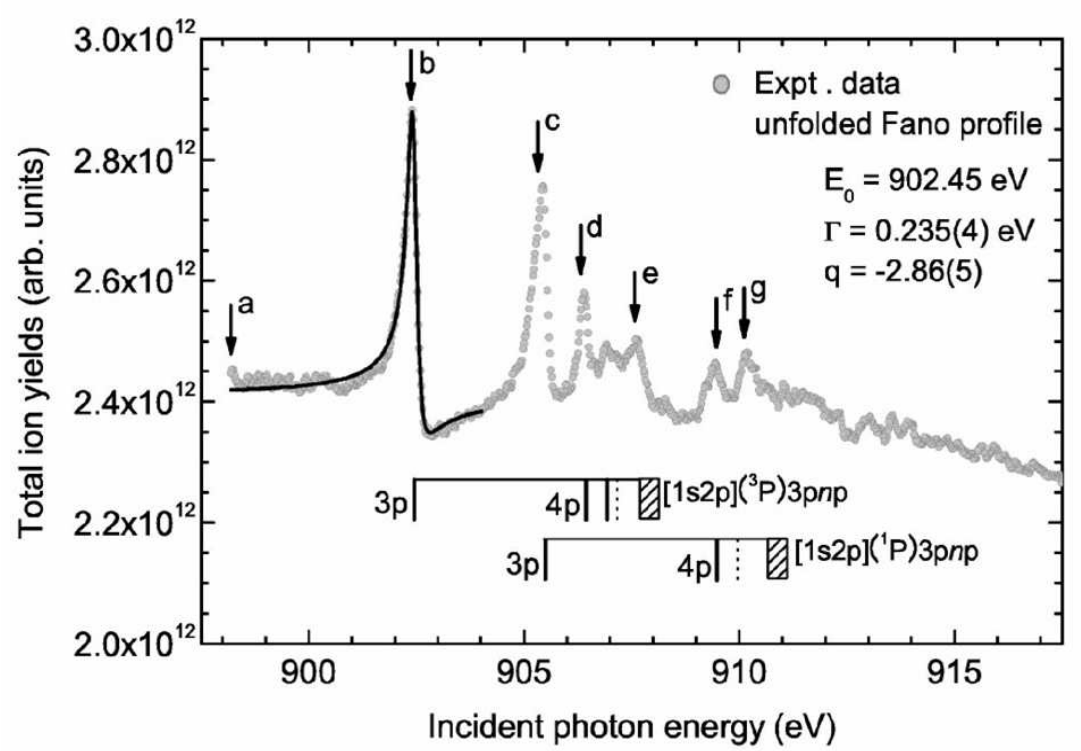


Fig. 9: *Double excitation spectrum of Ne with a Fano profile fit of the most intense line. The bars show the energies of the* $[1s2p](^{1,3}P)3pnp$ *series of transitions [33, 34].*

tiple Scattering of the photoelectrons also happens in larger Molecules and solids. This and the occurrence of multi-electron excitations are the reasons why the XANES region is extended to about 30 eV beyond the absorption edge. In the EXAFS range a single scattering event is assumed (see below).

For compounds, XANES shows characteristic features for different coordinations of the absorber atom. Increasing the oxidation state of the absorber shifts the position of the absorption edge in the XANES to higher energies. This is demonstrated in Fig. 11 for some Manganese compounds [38].

The edge shifts linear with the valence state of the atom. This is qualitatively easy to understand: The radial distribution of the valence-electron charges contributes to the screening of the nuclear charge also in the vicinity of the nucleus. The transfer of valence electrons to the ligands reduces this screening. As a consequence the core electrons get less screened and stronger attracted to the nucleus and it will cost more energy to bring them to an excited state.

Some times one also observes peaks in the pre-edge region close to the onset of the absorption edge. These could be due to quadrupole transitions ($s \rightarrow d$ at the K or L_1 edge or $p \rightarrow f$ at

the $L_{2,3}$ edges) to empty bound states. The selection rule for these to be allowed is $\Delta l = \pm 2$. However, quadrupole transitions are generally very weak in intensity. But, for transition metal compounds the pre-edge peaks are sometimes quite strong. They are then attributed to $p - d$ hybridization due to the interaction with ligands. In this case the transitions get dipole allowed due to orbital mixing.

Fig.12 shows as an example the K -edge absorption of oxygen coordinated chromium. First, we see the expected edge shift to higher energies when going from the 3+ to the 6+ valence state. The Cr^{3+} is 6-fold coordinated with an octahedral environment. With this coordination

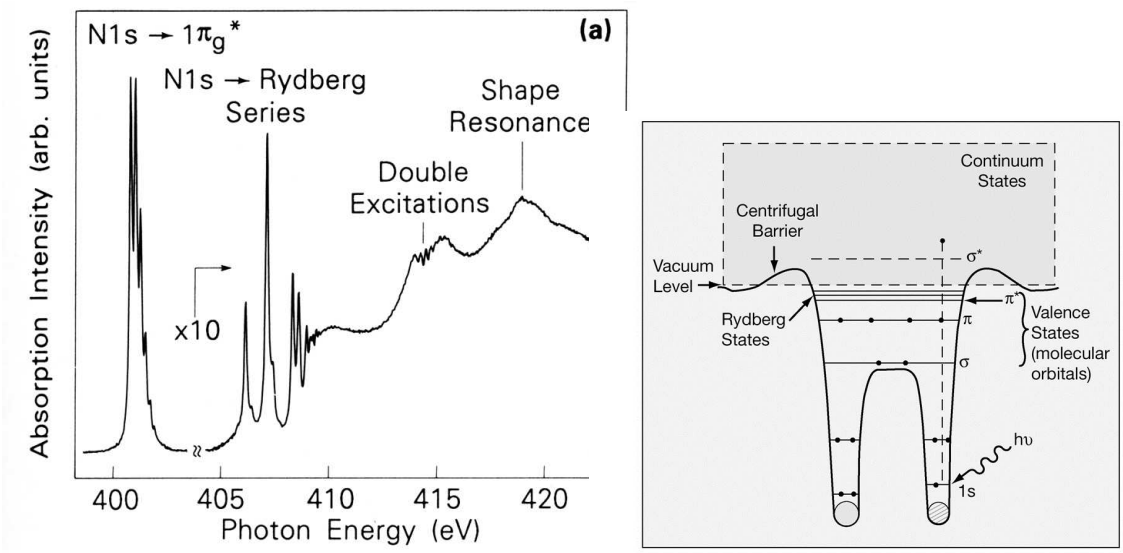


Fig. 10: *High resolution* K *-Shell photoabsortion spectrum of gaseous* N_2 *showing multi-electron excitations and a "shape resonance". The* $N1s \rightarrow \pi^*$ *resonance and the double excitation feature show vibrational fine structure. The Rydberg series merges in the ionization limit of* $\approx$ *410 eV [5, 35]. The figure on the right shows a scheme of a molecular potential including the centrifugal barrier (*$\propto l(l+1)/r^2$*) where here* l *is the angular momentum of the photoelectron with respect to the center of the molecule at* $r = 0$*. This potential barrier temporarily prevents the photoelectron from leaving the molecule and gives rise to the* σ^* *shape resonance above the ionization limit [36].*

in the ideal octahedral geometry - one has a center of inversion and $p - d$ mixing is not allowed[6]

The small features observed in the pre-edge region are therefore either due to quadrupole transitions or due to a weak $p - d$ mixing induced by a slight distortion of the octahedron. The Cr^{6+} in the chromate anion ($[CrO_4]^{2-}$) is fourfold coordinated with a tetrahedral environment.

[6] Orbitals can mix only if they belong to the same irreducible representation of the molecular point group. p-orbitals are antisymmetric ("ungerade") with respect to inversion and d-orbitals are symmetric ("gerade"), respectively. Therefore they do not mix as soon as the point group contains a center of inversion.

Neutral Cr has a $[Ar]3d^5 4s^1$ configuration. Thus, in its formal 6+ valence state the Cr has no d-electrons at all in its valence band. The strong p – d hybridization in the tetrahedral geometry will therefore provide a large density of empty valence states for the transition [7]. This example shows that the pre-edge region can give valuable information not only on the valence state but also on the coordination of the absorber atom.

A similar pre-edge peak has recently been reported for Fe(VI)-oxide in solution and is also interpreted as p – d hybridization with oxygen p-levels due to tetrahedral coordination. [40] The Fe(VI) has 8 empty d-orbitals and therefore also provides a high density of empty d-states which together with the hybridization also gives raise to a strong pre-edge peak.

According to Fermi's Golden Rule (eq. 6) the transition matrix element depends on the polarization of the X-rays. This can be exploited with anisotropic samples to determine bonding directions of the final state. The absorption cross section depends on molecular orientation. Therefore, XANES can be used to determine the orientation of molecules on surfaces [5, 41, 42].

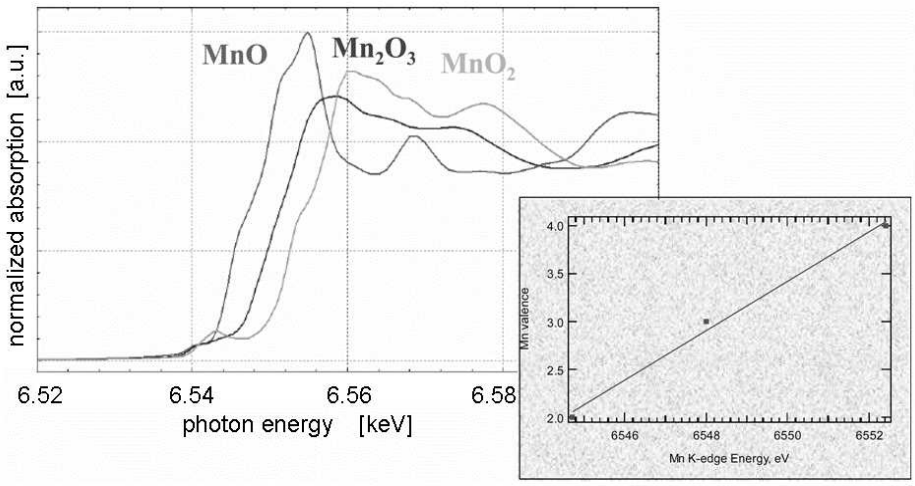


Fig. 11: *Shift of the Mn* K *-edge position for three different manganese oxides. As shown in the inset the position of the absorption edge for* Mn^{2+}, Mn^{3+}, *and* Mn^{4+} *compounds shifts linearly with the oxidation state of the absorber atom. This property can be used to identify the oxidation state of transition metal ions in unknown or complicated compounds like biomolecules [38].*

[7] Cr^{6+} compounds are very toxic. Therefore, this technique might help to identify them in critical cases.

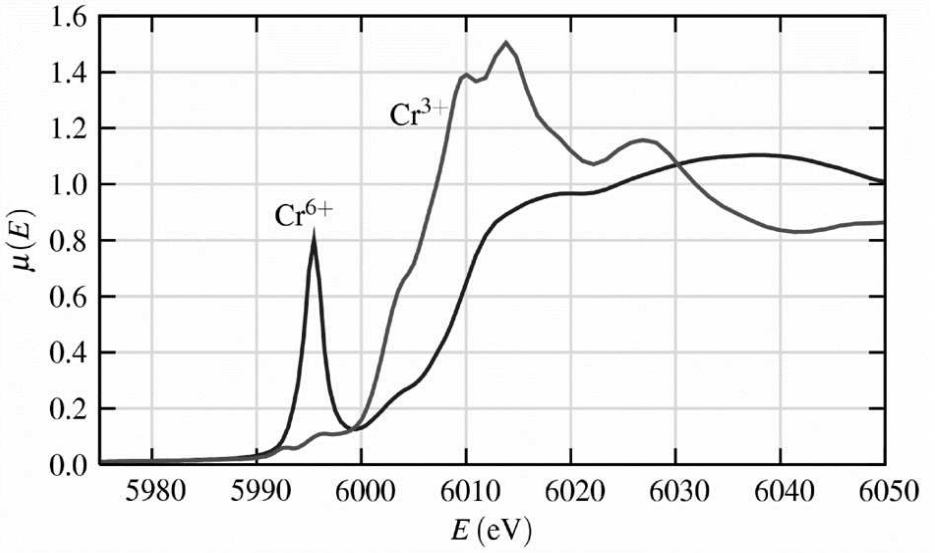


Fig. 12: *Cr K-edge XANES for Cr3+ and Cr6+ oxides. Here the strong pre-edge peak in the Cr6+ spectrum is a consequence of the tetrahedral symmetry causing considerable mixing of the empty d-electron orbitals with the p-states that the photo-electron will occupy [39].*

In conclusion, the XANES and the pre-edge structures tell us a lot about electronic structure. XANES is element specific, it probes the local unoccupied partial density of states in accordance with dipole selection rules. It is sensitive to the oxidation state and the coordination of the selected elemental species investigated and can be used to determine the orientation of adsorbate molecules. Difficulties arise, however, to interpret the spectral features in a quantitative way. This can easier be obtained with the EXAFS that we will discuss in the next chapter.

4 EXAFS (Extended X-ray Absorption Fine Structure)

About 30 eV beyond the absorption edge begins the EXAFS region [8,29,39,43-45]. In this range the photon energy is far above the ionization threshold and the photoelectron is treated as an outgoing spherical wave as schematically shown in Fig. 13.

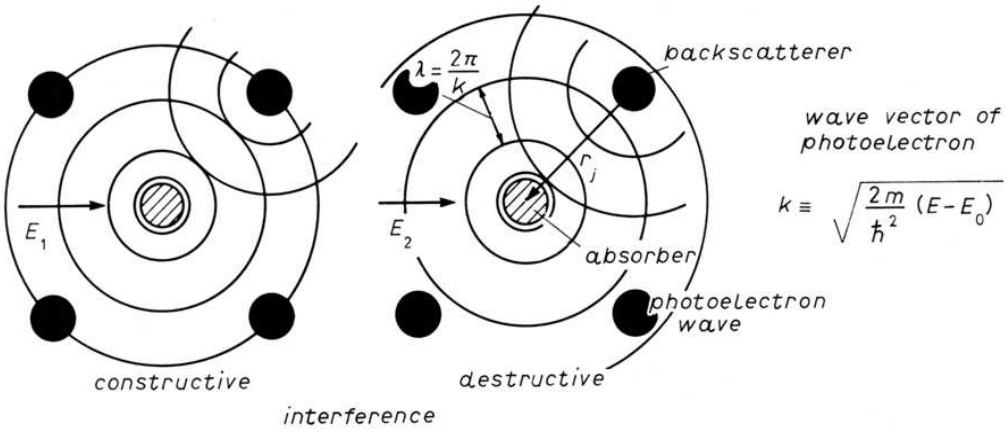


Fig. 13: *Scheme of the outgoing and backscattered electron wave leading to the EXAFS oscillations.* E *is the energy of the absorbed photon and* E_0 *the core level binding energy [8, 29].*

If there are other atoms surrounding the absorber atom they will scatter this wave. The outgoing wave and the scattered waves will interfere. The final state is the superposition of the outgoing and the scattered waves. For a constructive interference the signal will be enhanced leading to a maximum in the EXAFS oscillation. A destructive interference will lead to a minimum. The oscillations are absent in case of an isolated absorber atom. To quantify these ideas it is convenient to convert the energy scale into a wave number scale of the photoelectron. From the photoelectron's kinetic energy $E_{kin} = E - E_0 = p^2/2m = (nk)^2/2m$ we get

$$k = \sqrt{\frac{2m}{\hbar^2}(E - E_0)} \tag{8}$$

where E is the energy of the absorbed photon and E_0 is the energy of the ionization threshold. Then, using a plain wave approximation at the scattering atoms and allowing only for a single scattering event one can derive the following equation for the EXAFS oscillations [29,39,43]:

$$\chi(k) \equiv [\mu(E) - \mu_0(E)]/\mu_0(E) = \sum_j N_j F_j(k) D_j(k) e^{-2\sigma_j k^2} \frac{\sin\left(2kr_j + \phi_j(k)\right)}{kr_j^2} \tag{9}$$

Here $\mu_0(E)$ is the signal of the isolated absorber atom, which has been subtracted from the measured signal and to which the EXAFS oscillations are normalized. Of course, this function is not known but can be approximated by an extrapolation e.g. of the Victoreen function below and above the absorption edge [8]. The EXAFS oscillations are summed over all neighboring shells $\mathbf{j}$ of scatterers which are separated from the absorber by the same distance $\mathbf{r_j}$. N_j is the number of scatterers in one shell. The periodicity of the oscillations is determined by the sin function, where $2\mathbf{r_j}$ is the double distance between the absorber and a scatterer and φ_j is the total phase shift of one round trip of the wave. $F_j(k)$ is the backscattering amplitude of the neighboring atoms. The exponential is a Debye-Waller factor which describes the damping of the wave due to statistical disorder of the atoms and due to thermal motion [9]. Inelastic scattering of the photoelectron destroys its coherence. This is taken care of by the additional damping factor $D_j(k) = e^{-2r_j/\Lambda}$ where $\Lambda(k)$ is the mean free path of a photoelectron. The EXAFS formula has been derived under some simplifying assumptions: The photoelectron is treated as a plane wave at the positions of the neighbors. Multiple scattering contributions are not contained. These are largest for small k-values and for collinear arrangements of 3 atoms in a row. Moreover, the energy E_0 is not well defined. Therefore the EXAFS is often compared to a well characterized model compound from which better parameters can be deduced. Thereby it is hoped that the parameters are identical and that the uncertainties will cancel out. For larger k values respective shorter wavelength the dephasing increases and the oscillations will die away. Today computer codes are available which take corrections due to the curvature of the photoelectron wave as well as multiple scattering events into account.

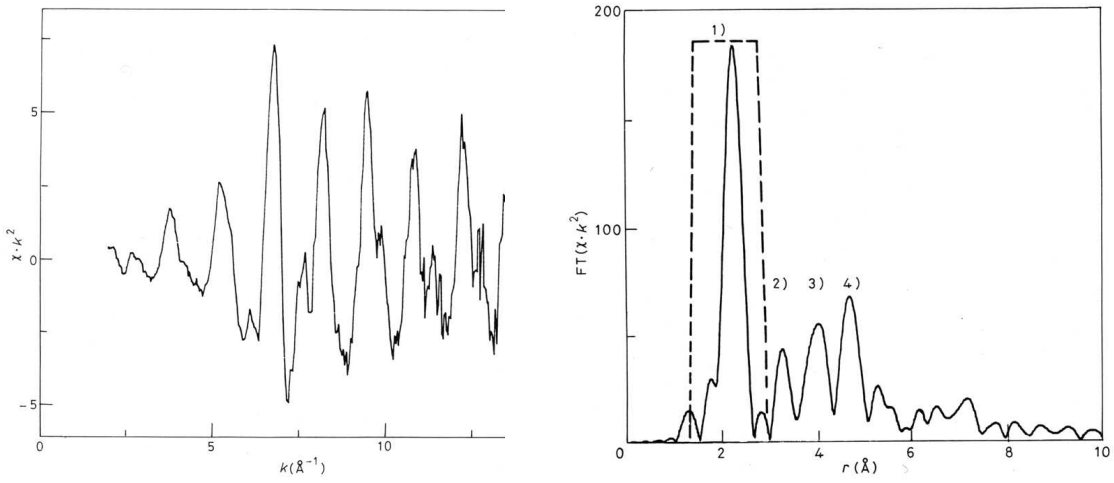


Fig. 14: *(a)* k^2 *weighted EXAFS* $\chi \cdot k^2$ *at the Ti* K *edge of a diluted NiTi alloy containing 2% Ti. The measurements were made at 77 K. (b) Fourier transform of the EXAFS. Four shells of neighbours can be identified. The first shell peak is covered by a filter function.*

[8] there are other ways to handle this [29,39]

[9] to reduce the latter effect it is advantageous to measure at low temperatures.

Fig. 14 shows as an example the EXAFS oscillations at the Ti K edge of a diluted NiTi alloy with 2 at. % Ti, measured at 77 K. Since the oscillations decay quickly they have been weighted by k^2. These oscillations are represented in k-space. To get real space information, that is to derive interatomic distances, we have to fourier transform the data. This is done in Fig. 14b. One clearly sees displayed the radii of the first four Ni coordination shells around a Ti atom. It is found that the 12 nearest neighbours are shifted outwards by (0.022 ± 0.005)Å. The atoms in the other coordination shells are at the Ni lattice sites.

In summary, the EXAFS is a superposition of sinusoidal waves that decay in intensity as the incident energy increases from the absorption edge. The sine waves result from the interaction of the ejected photoelectron wave with the surrounding atoms. The amplitude and phase depend on the local structure in the neighborhood of the excited atom. This structure is determined by fitting a theoretical spectrum to the experimental data. The procedure yields information on the distance of neighboring atoms, their identity, and number, and by the Debye-Waller factor the degree of disorder in the particular atomic shell.

5 XMCD: X-ray Magnetic Circular Dichroism

Dichroism denotes the property of anisotropic materials to change color when irradiated with light in different directions and/or of different polarization. It is associated with direction- and polarization- dependent differences in optical absorption and with this understanding can be extended to other spectral ranges, particularly X-rays. Circular dichroism refers to absorption differences for left and right circularly polarized light, respectively [10] , and if induced by the presence of a magnetization we speak of magnetic circular dichroism.

XMCD can be observed with ferro- and ferrimagnetic systems [11]. It was first observed at the K-edge of iron [46]. Yet, much stronger signals can be observed at the transition metal $L_{2,3}$ edges of 3d transition metals. These are of particular interest because the d-electrons are the primary carriers of the magnetism. Therefore, in this introduction the discussion will be limited to these cases.

Fig. 15a shows the absorption-cross sections at the $L_{2,3}$-edges of Fe, Co, Ni and Cu films [47]. The strong absorption peaks at the edges are often referred to as the "white lines" [12]. Their intensities decrease with increasing occupation of the d levels. For copper they are nearly absent, due to its almost completely filled d shell. To extract from these spectra the pure $p \rightarrow d$ transition intensity a two step background (due to transitions $p \rightarrow s$, p-derived delocalized band states) is subtracted as indicated in the spectrum of iron. The step function for the L_3 edge is twice as high as the one for the L_2 edge because of the different occupation numbers of the $2p_{3/2}$ and $2p_{1/2}$ levels, namely 4 and 2, respectively.

After background subtraction the integrated white-line intensities reflect the number of holes in the d band (Fig. 15b) [47-49]:

$$I(L_3) + I(L_2) = Cn_h = C(10 - n) \tag{10}$$

where n_h is the number of d holes in the electronic ground state and n is the number of occupied d states. This rule is called the "charge sum rule". Here, and in the following we will assume that anisotropies caused by anisotropic charge and spin distributions in the crystallographic unit cells

[10] For circularly polarized X- rays the electric field vector rotates about the direction of propagation. According to the classical definition of circular polarization the wave is called right circularly polarized if the light vector turns to the right for an observer looking into the beam and left circularly polarized if the light vector turns left. However, in X-ray spectroscopy not all investigators cling to this definition. Therefore it might be more adequate to refer to the light helicity which is positive if the photon angular momentum points into the propagation direction of the light (σ^+ light; left circularly polarized according to the above rule)) and negative, respectively, if the photon angular momentum points in the opposite direction (σ^- light). X-ray natural circular dichroism (XNCD) has been observed in the uniaxial gyrotropic single crystal $LiIO_3$. [50]

[11] It has also been studied with paramagnetic samples. This requires low temperatures and strong magnetic fields of several Tesla. Here the net spin orientation is obtained by the Zeeman interaction. Some applications to bio-inorganic molecules and molecular magnets are summarized by Funk et al. [51]. We will not discuss this any further.

[12] This nomenclature stems from the early days of X-ray spectroscopy when the absorption was still measured with photographic films. The films were not exposed at the position of the strong absorption lines and therefore appeared to be white.

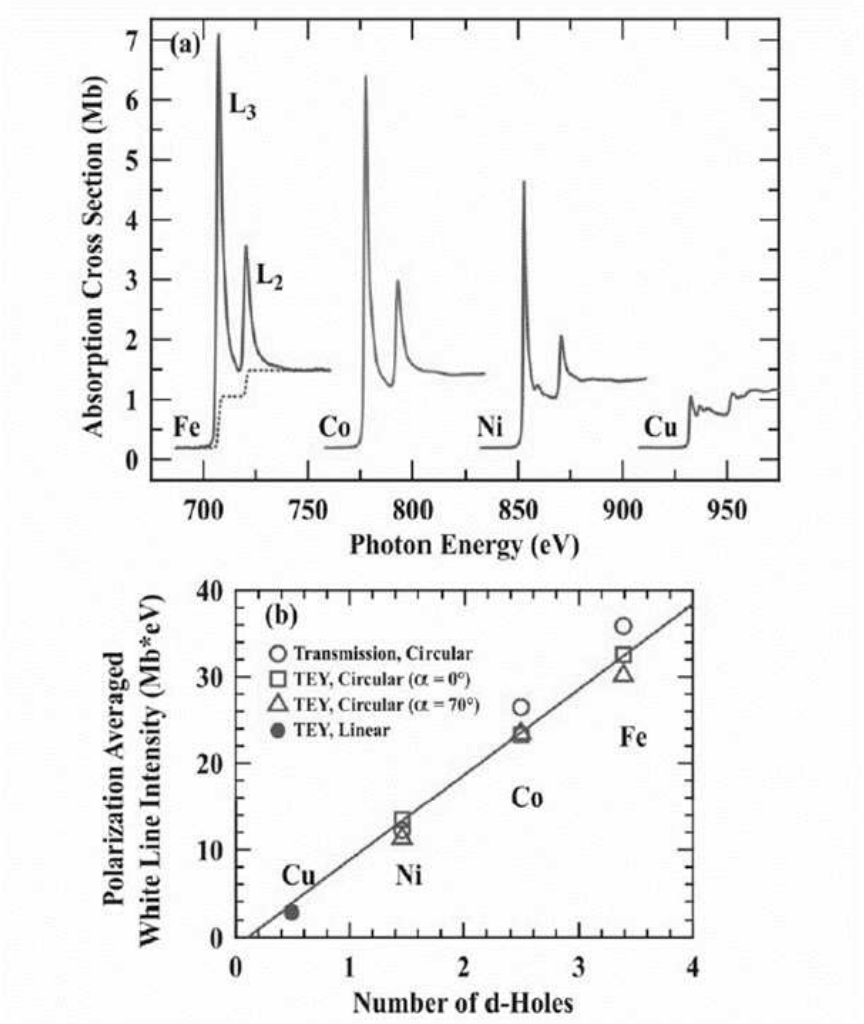


Fig. 15: *L edge X-ray absorption spectra for Fe, Co, Ni, and Cu metal. (b) White-line intensity, determined after subtraction of a double step function [52], shown dashed in the Fe spectrum in (a), versus calculated number of* d *holes [47].*

are averaged out. Experimentally this can be achieved by averaging over three measurements performed in three orthogonal directions [49]. In a tight binding approximation the constant C is calculated to be: $C = \frac{2}{15} A\Re$ where $A = 4\pi^2 n\omega\alpha_f$ and $\Re$ is the radial transition matrix element [49, cf. appendix].

$\hbar\omega$ is the photon energy and $\alpha_f = \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}$ is Sommerfeld's fine structure constant.

Thus XAS probes the unoccupied density of states (DOS) above the Fermi level.
If, in addition to relation (10), we could distinguish between spin-up and spin-down DOS we were able to measure the magnetic moments of the samples with element selectivity on an atomic level. Therefor the excitation of core electrons has to be spin selective. This can

indeed be achieved by using circularly polarized light. It requires the presence of two basic inter- actions, namely spin-orbit coupling and an exchange splitting of the electronic valence states. With the aid of Fig. 16 we will explain the role of the spin-orbit coupling. In Fig. 16 we study the transitions from 2p initial states to 3d states when we excite the 2p- states with σ^+ light. The transitions are energetically separated by spin-orbit interaction.
The states can be described as linear combinations of $|m_l, m_s>$ kets and it is sufficient to study transitions between those. Transitions from the $2p_3$ and $2p_1$ initial states are treated separately.

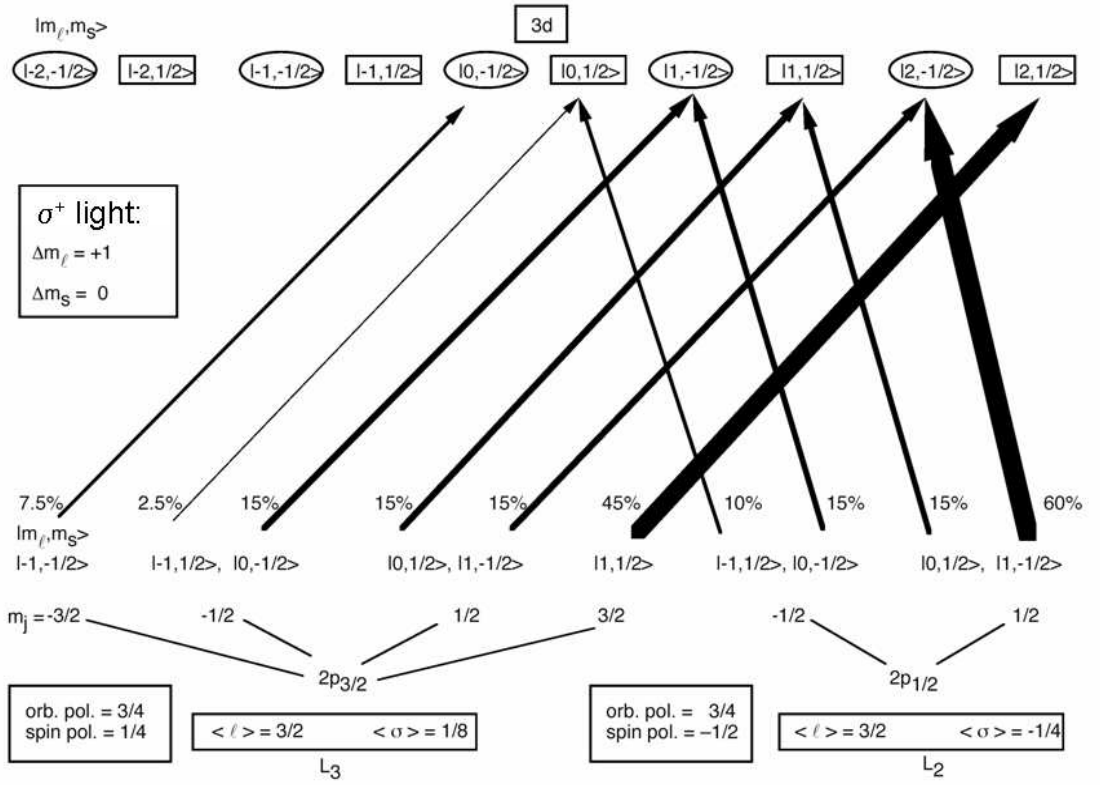


Fig. 16: *Scheme of possible* 2p → 3d *transitions when excited with* σ^+ *light [53]. The upper row shows the ten* 3d *final state levels sorted by their* m_l *and* m_s *quantum numbers. Spin-orbit splitting is neglected for the* 3d *states. They are labeled by rectangles (positive* m_s *) and ellipses (negative* m_s*). It is assumed that all* d *levels are unoccupied prior to the transition (there are no restrictions due to the Pauli-exclusion principle). Below, the spin-orbit split 2p levels are given. The magnetic sublevels are indexed by their magnetic quantum number* m_j .
The upper line shows uncoupled m_l,m_s *states that contribute to the respective* m_j *states according to the Clebsch-Gordon series (see appendix). Allowed transitions - represented by arrows - must obey dipole selection rules (for* σ^+ *light:* $\Delta m_s = 0$ *and* $\Delta m_l = +1$*). The line thickness corresponds to the transition strength. Separately for each spin-orbit split component the transition strength is also given in % as calculated in the appendix. For the* $2p_{\frac{1}{2}}$ *level 75 %are spin down transitions and 25 % spin up. This results in a net spin polarization of* $\frac{-75+25}{75+25} = -\frac{1}{2} = -50\%$

The spin-orbit splitting of the d states, which is much smaller than that of the 2p states is neglected. For σ^+ light the transitions have to obey the dipole selection rules with $\Delta m_l = +1$ and $\Delta m_s = 0$. Here it is assumed that all the d-states are unoccupied and can be reached (no restrictions due to the Pauli principle). The width of the arrows corresponds to the calculated transition probabilities (In the appendix it is shown how this result is derived from calculations of transition matrix elements). The transition strengths are also given in %. If for the $2p_{1/2}$ initial state we add all the transitions with identical spin we find a total probability of 75 % for a transition with $m_s = -1/2$ (spin down) while transitions with $m_s = +1/2$ (spin up) contribute only 25 %. As a result we end up with a spin polarization of 50 % (spin down) in the final d-states if only $2p_{1/2}$ states are excited. The spin expectation value is then $\langle\sigma\rangle = -1/4$.[13] Similarly, if only $2p_{3/2}$ initial states are excited we get 37.5 % spin down and 62.5 % spin up electrons in the excited states. This results in a net spin polarization of 25 % (spin up) electrons and a spin expectation value $\langle\sigma\rangle = -1/8$. Note, that the spin polarization is opposite in the two cases. Since we have twice as many $2p_{3/2}$ states as $2p_{1/2}$ states the absorption strength of the $2p_{3/2}$ states is also doubled. Thus, if there is no spin-orbit splitting and all the states are excited simultaneously the net spin polarization in the excited state will be zero, as expected. One also observes an orbital polarization in the excited states. The orbital polarization is 75 % for both excitation pathways where $m_l = 2$ is fully polarized, $m_l = 1$ half-polarized and $m_l = 0$ non-polarized. When the polarization of the exciting photon is reversed ($\sigma-$ light), the polarization of the excited electrons is reversed as well. Thus, if the d states are unoccupied there is no difference in the absorption for both circular polarization directions of the light. The same is true if the d-states are partially occupied with electrons but the same number of electron spins points towards the beam as in the opposite direction.

The amazing fact, that spin polarized electrons can be generated by photoexcitation of unpolarized atoms with circularly polarized light if the states can be separated by spin-orbit splitting was predicted by U. Fano in 1969 [54]. The effect is therefore called Fano effect.[14] It is now widely used to generate spinpolarized electrons from GaAs sources [55-58].

Above, it was stated that, besides spin-orbit coupling, we need exchange interaction to observe the circular dichroism. In the valence band of a ferromagnet (as well as in a ferrimagnet) the exchange interaction leads to different densities of states for majority and minority spins, $\rho\uparrow$ and $\rho\downarrow$. Then there are different numbers of d holes available for σ^+ and σ^- light. This destroys the absorption symmetry. The XMCD signal is just the difference of the absorption for these two cases.

Fig. 17 (left) shows the magnetic circular dichroism spectrum of an ultra thin Co film on Cu(001)[53,60,61]. In the upper panel, the absorption spectra in the photon energy range of the Co 2p → 3d transitions are shown for σ^+ and σ^- light. For a fixed sample magnetization the absorption at the L_3 peak (corresponding to $2p_{3/2} \rightarrow 3d$ transitions) is higher for σ^+ polarization (continuous line) compared to $\sigma-$ polarization (dashed line), while the situation is opposite at the L_2 peak ($2p_{1/2} \rightarrow 3d$ transitions). The bottom panel displays the dichroism

[13] Note, that $\langle\sigma\rangle = 1/2$ means 100 % down spin polarization.

[14] not to be mixed-up with the earlier mentioned Fano resonance.

spectrum, which shows different sign at the two Co absorption edges. The signal scales as the inner product q · M where q is the propagation direction of the circularly polarized photons and M is the sample magnetization[15].

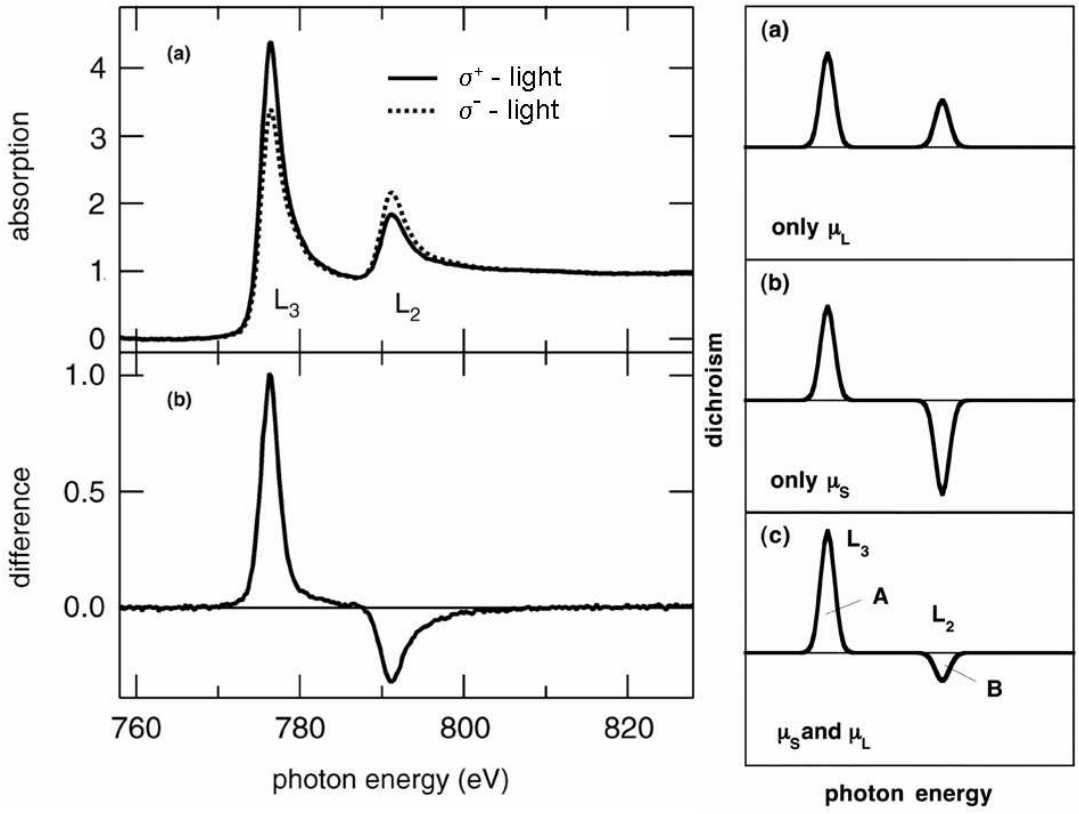


Fig. 17: *Left: XMCD spectrum of a thin Co film on Cu(001). (a): Absorption spectra at the Co* $L_{2,3}$ *edges for* σ^+ *(continuous line) and* σ^- *polarization (dotted line) of the exciting X-rays. The magnetic dichroism is visible as a difference in the absorption strength.(b): Difference of the two spectra (XMCD spectrum). The sign of the dichroism is opposite at the* L_3 *and* L_2 *edges. This is due to the inverse spin polarization of the* $2p_{3/2} \rightarrow 3d$ *and* $2p_{1/2} \rightarrow 3d$ *transitions. Right: Scheme to illustrate the separate orbital and spin contributions to the XMCD spectrum. (a) hypothetical XMCD spectrum obtained from orbital magnetic moments alone, (b) hypothetical spectrum with contributions of the spin magnetic moments alone, (c) XMCD difference spectrum of a sample with both spin and orbital magnetic moments. [60,61].*

In the appendix we have shown how within a simple model we can calculate the XMCD signal starting from a magnetic ground state with a given d orbital occupation. This procedure can principally be inverted. Important sum rules have been derived to extract from the experimental

[15] This rule holds for absorption and total yield measurements. Deviations may occur when partial yield detection is used [62]. Therefore, with electron detection modes special care has to be taken.

spectra the ground state expectation values of the orbital and spin angular momenta, separately [63-68, see also 48,49,52,69]. The "orbital sum rule" [63,65,66] relates the average orbital moment per atom with the peak intensities according to

$$\mu_l = -\frac{2}{3C}(A+B)\mu_B$$

The constant C is the same as the one for the charge sum rule given above (10). The notation for the A and B is given in fig.17 (right panel). Correspondingly, the "spin sum rule" [64] links the dichroism intensities with the size of the spin moment per atom according to

$$\mu_s = -\frac{1}{C}(A-2B)\mu_B$$

Note, that the area of B is negative. Therefore in the derivation of μ_l the difference of the two areas has to be taken, which will lead to a considerable error when both areas come close to each other.

To make the sum rules plausible we follow the line drawn by Kuch [60,61] (Fig 17 right panel). From Fig. 16 and the calculation given in the appendix we know that the orbital contribution to the $2p_{3/2} \rightarrow 3d$ transition is twice that of the $2p_{1/2} \rightarrow 3d$ transition. Upon reversal of the light helicity from σ^+ to σ^- spin and orbital polarizations will both change sign, i.e. in a difference spectrum they will both add in the same sense. Thus a hypothetical spectrum with only orbital admixture would give a spectrum as plotted in Fig. 17(a)(right panel) where peak A has the double size of peak B. If, on the other hand, we consider a spectrum with only spin contribution the peaks A and B have opposite sign as shown in the middle part of the figure. The higher spin polarization of transitions at the L_2 edge compensates the lower electron occupation of the $2p_{1/2}$ level leading to opposite peaks of equal size. This is also in agreement with our calculation and Fig. 16. A realistic spectrum with both, spin and orbital contributions, would be a linear combination of these two (Fig 17(c)). Now we see, that addition of peaks A and B will eliminate the spin part and retain the orbital part only. If we take (A − 2B) we will eliminate the orbital part and retain only the spin part. A more rigorous calculation gives the results reported in the sum rules above.

So far, it was assumed that anisotropies are averaged out. This is acceptable for bulk transition metals. More general forms of the sum rules are discussed elsewhere [48,52,69,70]. Particularly, for thin films the spin magnetic moment has to be supplemented by a magnetic dipole term which can be of the same order of magnitude as the orbital magnetic moment.

The sum rules are derived under some simplifying assumptions. For instance the electronic transitions are considered to take place between states of free atoms. Many particle effects are neglected. Nevertheless they are considered to be correct within $\approx \pm 10$ %. The reader is refered to refs. [48,49,52,69,70] for more detailed discussions. A very readable article with a user friendly discussion of the general sum rules is given in [69].

It is one of the major achievements of XMCD that it gives experimental access not only to the spin magnetic moment but also to the orbital magnetic moment which is closely related to the

magnetic anisotropy energy [71]. The orbital moment is small in solids but its contribution to the total magnetic moment/atom increases as the sample dimensions become smaller from thin films to nanowires to clusters [72-79].
One of the major applications of XMCD today is the study of magnetic domains on a microscopic level in a photoelectron emission microscope (PEEM). Due to its element selectivity and its relatively large penetration depths it offers the unique possibility not only to study the magnetic structure of multilayers layer by layer but also the coupling between layers that are separated by nonmagnetic spacer layers. Such structures are of great technological importance in todays magnetic storage devices and for spintronic applications [16].

6 XMLD: X-ray Magnetic Linear Dichroism

Besides the XMCD there exists a magnetic linear dichroism, which can be observed with linearly polarized X-rays with the polarization direction aligned parallel and perpendicular to the magnetic field, respectively [80,81].
The application of a magnetic field to a medium (internally or externally) breaks its symmetry. According to Curie's principle [82] [16] the medium gets uniaxially distorted, e.g. a cubic crystal exposed to a magnetic field will show a uniaxial symmetry. This can be probed by linearly polarized light and will be invariant with respect to a reversal of the magnetic field. The anisotropy will therefore scale as (M^2) to first order [80]. The largest anisotropy will occur when the propagation direction of the light is perpendicular to the magnetic field and the polarization direction is changed from parallel to the magnetic field to perpendicular. This is the basis of XMLD. It is one of the major advances of XMLD that can be applied to antiferromagnetic materials as well. Such materials are used in magneto-electronics to pin the magnetization of ferromagnetic layers by an exchange biasing interaction.

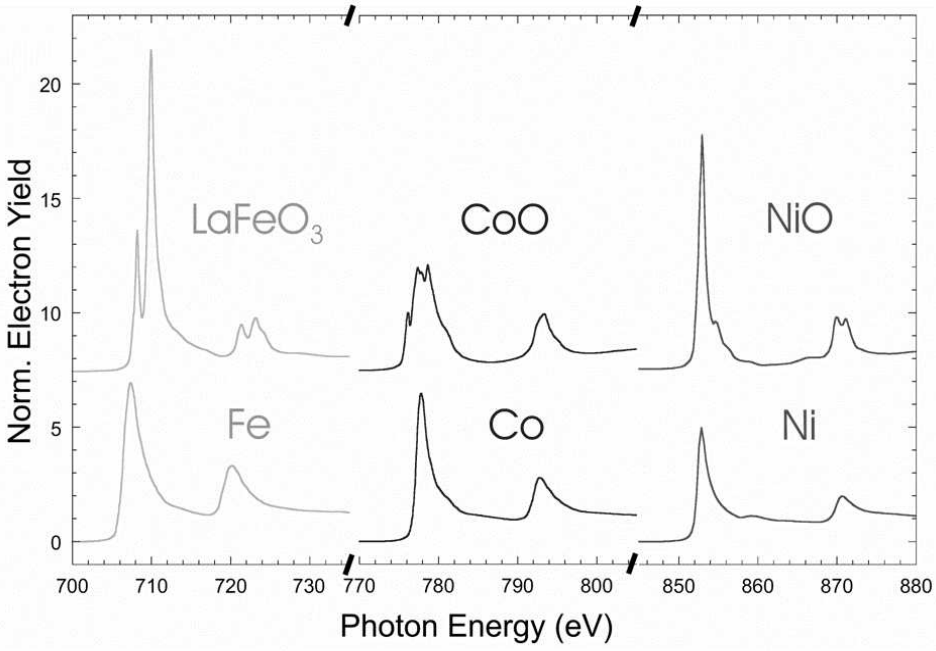


Fig. 18: *L edge X-ray absorption spectra for Fe, Co, Ni, compared to those of some antiferromagnetic oxides.* 2p *core electrons are excited to empty* d *states in all the spectra. The oxide spectra show more detailed fine structure due to multiplet splitting and crystal field effects [84].*

Fig. 18 compares X-ray absorption spectra of Fe, Co, and Ni with those of some antiferro-

[16] The Curie principle states that the symmetry group of the system with an influence (here the magnetic field) G_m contains a common subgroup of the symmetry group of the system without this influence G_s and the symmetry group of the influence G_i alone: $G_m \supseteq G_s \cap G_i$.

magnetic oxides. The spectra of the antiferromagnets have a richer structure which cannot be interpreted in terms of a simple single particle picture. The structures occur due to multiplet splitting and crystal field effects [83]. Fig. 19 demonstrates the XMLD at the L_2 edge of a NiO (100) thin film when the light polarization is changed from parallel to the magnetic field (solid line) to perpendicular (dashed line).

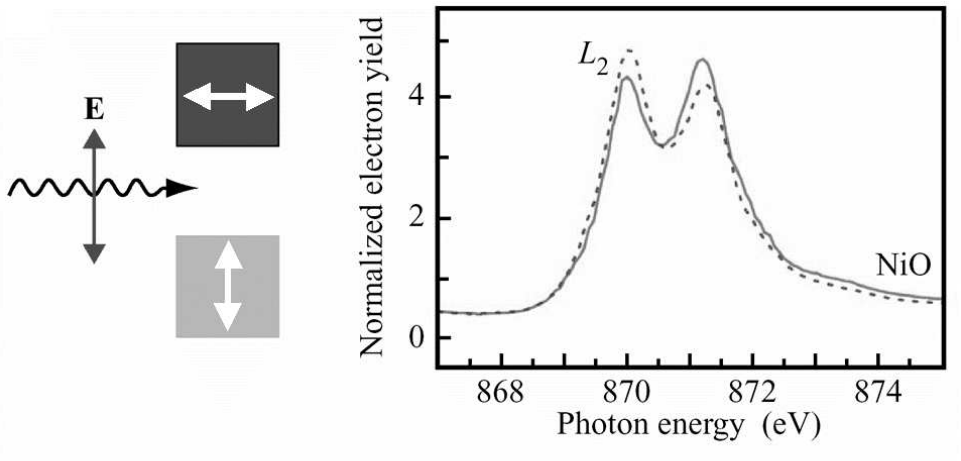


Fig. 19: *XMLD spectrum of a NiO film at the* L_2 *edge. The spectrum was recorded by electron yield detection for 45 nm thick NiO (100) film grown on MgO (100). The polarization of the X-rays was changed from parallel to the magnetic film (solid line) to perpendicular (dashed line) [85,86].*

In NiO successive (111) lattice planes have opposite spin alignment. Thus ”uncompensated” surfaces can be generated at the surface of a single crystal. Within the surface the magnetic moment has three equivalent possibilities to orient, namely in $[\bar{2}11]$, $[1\bar{2}1]$ or $[11\bar{2}]$ directions. Together with the four equivalent (111) surfaces this results in 12 possible spin orientations, and, thus in 12 different possibilities to form antiferromagnetic domains. To proof that the anisotropy is due to the antiferromagnetic alignment and not due to another anisotropic effect one can heat the sample to its Neél temperature (520 K for NiO) where the magnetic contribution will disappear.
XMLD can also be used for domain imaging by means of photoemission electron microscopy, PEEM [16]. This is of coarse particularly useful for antiferromagnets.

7 Conclusion

X-ray absortion spectroscopy is a local element-specific probe of the electronic and geometric structure of materials. It can likewise be applied to ordered and disordered samples. It can be used to study subsurface structures and does not rely on UHV conditions. The latter is particularly useful for catalytic interfaces and biological samples. At grazing incidence X-ray absorption gets surface sensitive with probe depths of 2-5 nm. In the XANES (NEXAFS) range it probes the projected electronic density of empty states (local partial density of States LPDOS) and the valence of absorber atoms. EXAFS probes the coordination number and the interatomic distance.
X-ray absorption spectroscopy probes transitions between core levels and empty valence states. It is not only element specific but also sensitive to the chemical environment. It is particularly useful for investigations of magnetic properties. XMCD measures the size of the magnetization per atom and its angle relative to the circular x-ray polarization vector. Sum rules exist for the number of d holes, the spin and the orbital moment. XMCD is used for the study of ferromagnets. XMLD measures the magnetic moment and the angle between the spin axis and the linear x-ray polarization vector. XMLD is used for the study of ferro- and antiferromagnets. In connection with photoemission electron microscopy (PEEM) X-ray absorption allows to visualize magnetic domains with a resolution of better than 100 nm. Dynamical processes can as well be studied. In connection with the free electron laser it will become possible to study transient electronic and dynamic processes on a femtosecond time scale.

8 Appendix

To obtain the transition rates in Fig.16 we have to evaluate the transition matrix element in Fermi's golden rule (equ.(6)). The following approach was first used by J. L. Erskine and E. A. Stern [87, see also:19,49,59,69]. For circularly polarized light propagating in the positive z-direction the normalized polarization vector for positive and negative helicity of the electric field is

$$\epsilon^{\pm} = \frac{e_x \pm i e_y}{\sqrt{2}}$$

where e_x and e_y are unit vectors pointing in the direction of the coordinate axes. The position vector is:

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z .$$

Thus, the polarization-dependent dipolar operator is given as:

$$\epsilon^{\pm} \cdot r = \frac{x \pm iy}{\sqrt{2}}$$

Introducing spherical harmonics

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}}\frac{z}{r} = \sqrt{\frac{3}{4\pi}}\cos\theta \quad and \quad Y_{1,\pm1} = \sqrt{\frac{3}{8\pi}}\frac{x \pm iy}{r} = \sqrt{\frac{3}{8\pi}}\sin\theta\, e^{\pm i\varphi}$$

where (r, θ,) are spherical coordinates. With:

$$x = r\sin\theta\cos\varphi = -\sqrt{2\pi/3}\, r\left(Y_{1,1} - Y_{1,-1}\right); \quad y = r\sin\theta\sin\varphi = i\sqrt{2\pi/3}\, r\left(Y_{1,1} + Y_{1,-1}\right)$$

$$z = r\cos\theta = \sqrt{4\pi/3}\, rY_{1,0}$$

we get the polarization dependent dipolar operator:

$$\epsilon^{\pm} \cdot r = \frac{x \pm iy}{\sqrt{2}} = \mp r\sqrt{\frac{4\pi}{3}}Y_{1,\pm1} = \mp rC^{(1)}_{\pm1}$$

$C^{(l)}_m = \sqrt{4\pi/2l+1}\, Y_{l,m}(\theta,\varphi)$ is Racah's spherical Tensor operator [88] and is introduced here to simply eliminate the numerical factor.
Dividing the transition rate (Fermi's golden rule) by the photon flux gives the absorption cross- section in the dipole approximation:

$$\sigma^{\pm} = 4\pi^2\hbar\omega\alpha \sum_{f,i} \left|\left\langle f\left|\mp r C_{\pm 1}^{(1)}\right|i\right\rangle\right|^2 \delta(E_f - E_i - \hbar\omega)$$

where α is Sommerfeld's fine-structure constant. The asymmetry of the circular dichroic signal is then given as $A = \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-}$

According to dipole selection rules the transition matrix elements are non-zero only if $\Delta \mathbf{l} = l_f - l_i = \pm 1$ and $\Delta m_s = 0$ and $\Delta m_l = +1$ for left circular polarization (positiv helicity) and $\Delta m_l = -1$ for right circular polarization (negativ helicity). The absorption intensity is given as:

$$I \propto \sum_{f,i} \left|\langle \psi_f | x \pm iy | \psi_i \rangle\right|^2 \delta(E_f - E_i - \hbar\omega)$$

Because of the spin-orbit coupling the six 2p states are split (at least in the final state configuration (see above)). These are eigenstates of total angular momentum j^2 and j_z with quantum numbers $\mathbf{j}$ and m_j . The new states $|j,m_j\rangle$ are then related to the uncoupled states $|l,m_l\rangle$ and $|s,m_s\rangle$ by the relation:

$$|j,m_j\rangle = |ls;j,m_j\rangle = \sum_{m_l m_s} \langle ls, m_l m_s | j,m_j\rangle |l,m_l\rangle |s,m_s\rangle = \sum_{m_l m_s} \langle ls, m_l m_s | j,m_j\rangle |l,m_l; s,m_s\rangle$$

where $\mathbf{j} = l \pm s$ and $m_j = m_l + m_s$. $\langle ls, m_l m_s | j,m_j\rangle$ are the Clebsch-Gordon coefficients. For the level with maximum values of the quantum numbers $\mathbf{j}$ and m_j the Clebsch-Gordon coefficient is 1. The others can be calculated with the aid of ladder operators or simply by use of recursion formulas deduced with those [89]. Many people use to prefer the Wigner "3**j**" symbol instead of the Clebsch-Gordon coefficients because of its higher symmetry. Both coefficients are related by:

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \frac{(-)^{j_1 - j_2 + M}}{\sqrt{2J+1}} \langle j_1 j_2 m_1 m_2 | J,M\rangle$$

If it is assumed that the overlap of the radial part of the wavefunction is constant and the same for the L_2 and the L_3 edge, and since the dipole operator does not act on the spin part of the wavefunction the polarization dependence of the transitions is entirely contained in the angular parts of the matrix elements: $\left\langle l = 2, m_l \pm 1 \left| \mp C_{\pm 1}^{(1)} \right| l = 1, m_l \right\rangle$. Analytical solutions to this problem are known [90]. The calculated values with the wave functions given above are summarized in the following table:

J	m_j	m_l	m_s	$30M^2$	M^2	$(CG)^2$	I_{calc}	I%
$^1/_2$	$-^1/_2$	-1	$^1/_2$	2	$^1/_{15}$	$^2/_3$	$^2/_{45}$	10
$^1/_2$	$-^1/_2$	0	$-^1/_2$	6	$^1/_5$	$^1/_3$	$^1/_{15}$	15
$^1/_2$	$^1/_2$	0	$^1/_2$	6	$^1/_5$	$^1/_3$	$^1/_{15}$	15
$^1/_2$	$^1/_2$	1	$-^1/_2$	12	$^2/_5$	$^2/_3$	$^4/_{15}$	60
$^3/_2$	$-^3/_2$	-1	$-^1/_2$	2	$^1/_{15}$	1	$^1/_{15}$	7,5
$^3/_2$	$-^1/_2$	-1	$^1/_2$	2	$^1/_{15}$	$^1/_3$	$^1/_{45}$	2,5
$^3/_2$	$-^1/_2$	0	$-^1/_2$	6	$^1/_5$	$^2/_3$	$^2/_{15}$	15
$^3/_2$	$^1/_2$	0	$^1/_2$	6	$^1/_5$	$^2/_3$	$^2/_{15}$	15
$^3/_2$	$^1/_2$	1	$-^1/_2$	12	$^2/_5$	$^1/_3$	$^2/_{15}$	15
$^3/_2$	$^3/_2$	1	$^1/_2$	12	$^2/_5$	1	$^2/_5$	45

where M^2 is the square of the respective matrix element and $(C G)^2$ the square of the corresponding Clebsch Gordan coefficient. $I_{calc} = M^2 (C G)^2$ is the relative contribution to the absorption intensity (also given in %). These data reproduce the values as given in Fig.16.
So far, we have assumed that all d states are empty and accessible by transitions. This is not the case for a magnetic material, but we can now within this simple model sort out all occupied d states and calculate the XMCD signal. (Transitions to the occupied states are forbidden by the Pauli exclusion principle). For the simple case that all $m_s = -^1/_2$ states are occupied (half filled d band: L = 0) at the L_2 edge we get a 25 % σ_+ transition intensity:

$$I_{\sigma^+} = \left(\frac{1}{15} + \frac{2}{45}\right)\Re^2 = \frac{1}{9}\Re^2$$

and for σ_- polarized light we get a 75% transition intensity

$$I_{\sigma^-} = \left(\frac{1}{15} + \frac{4}{15}\right)\Re^2 = \frac{1}{3}\Re^2$$

Therefore, for this ground state occupation the dichroism at the L_2 edge is:

$$\Delta I_{L_2} = I_{\sigma^+} - I_{\sigma^-} = -\frac{2}{9}\Re^2$$

Similarly, at the L_3 edge we get:

$$I_{\sigma^+} = \left(\frac{1}{15} + \frac{2}{45}\right)\Re^2 = \frac{1}{9}\Re^2$$

and

$$I_{\sigma^-} = \left(\frac{1}{15} + \frac{4}{15}\right)\Re^2 = \frac{1}{3}\Re^2$$

The dichroism at the L_3 edge is:

$$\Delta I_{L_2} = I_{\sigma^+} - I_{\sigma^-} = \frac{2}{9}\Re^2$$

that is, in this simple case the dichroic signal at the L_3 and L_2 edges has equal magnitude but opposite sign as expected for a spin only signal. (The orbital momentum for a half filled band is zero.) In this example we have from a given ground state occupation deduced the corresponding XMCD- signal. The sum rules discussed in chapter 5 invert this procedure and deduce the ground state occupation from the observed XMCD signal.

References

[1] S. Cramm, IFF Spring School Vol. 38, 2007

[2] C.S. Fadley, IFF Spring School Vol. 38, 2007

[3] B.K. Agarwal, *X-ray Spectroscopy, An Introduction*, Springer Verlag Berlin, Heidelberg, New York (1979)

[4] D.C. Koningsberger and R. Prins(Ed.), *X-ray Absorption - Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*, John Wiley & Sons, New York (1988)

[5] J. Stöhr, *NEXAFS Spectroscopy*, Vol. 25 of Springer Series in Surface Science, Springer Verlag, Berlin, (1996)

[6] A.S. Schlachter and F.J. Wuilleumier (Ed.), *New Directions in Research with Third-Generation Soft X-Ray Synchrotron Radiation Sources*, Kluwer Academic Publishers, Dordrecht, Boston, London (1994)

[7] F. Hippert, E. Geissler, J.L. Hodeau, E.Lelievre, J.R. Regnard (Ed.), *Neutron and X-ray Spectroscopy*, Springer Dordrecht (2006).

[8] B. Lengeler in: *Synchrotronstrahlung zur Erforschung kondensierter Materie*, 23. IFF-Ferienkurs, Forschungszentrum Jülich (1992), ISBN 3-89336-088-3 (in German)

[9] U. Döbler, in: *Synchrotronstrahlung zur Erforschung kondensierter Materie*, 23. IFF-Ferienkurs, Forschungszentrum Jülich (1992), ISBN 3-89336-088-3 (in German)

[10] W. Zinn, in: *Synchrotronstrahlung zur Erforschung kondensierter Materie*, 23. IFF-Ferienkurs, Forschungszentrum Jülich (1992), ISBN 3-89336-088-3 (in German)

[11] G. Schütz,and P. Fischer, in: *Magnetismus von Festkörpern und Grenzflächen*, 24. IFF-Ferienkurs, Forschungszentrum Jülich (1993), ISBN 3-89336-110-3 (in German)

[12] H. Ebert and Ya. Perlov, in: *Magnetische Schichtsysteme*, 30. IFF-Ferienkurs, Forschungszentrum Jülich (1999), ISBN 3-89336-235-5 (in German)

[13] S. Eisebitt, in: in: *Magnetische Schichtsysteme*, 30. IFF-Ferienkurs, Forschungszentrum Jülich (1999), ISBN 3-89336-235-5 (in German)

[14] F.U. Hillebrecht, in: *Magnetism goes Nano*, 36. IFF-Ferienkurs, Forschungszentrum Jülich (2005), ISBN 3-89336-381-5

[15] R. Stumm von Bordwehr, Ann. Phys. Fr. **14**, 377(1989) [16] see contribution of C. M. Schneider to this Spring School

[17] J. A. Victoreen, J. Appl. Phys. **14**, 95 (1943); ibid. **19**, 855(1948) [18] W. Kossel,Phys.Z. **18**, 240 (1917)

[19] J. Stöhr and Y. Wu, in [6]

[20] G.K. Woodgate, *Elementary Atomic Structure*; Mc Graw Hill, London (1970), p.36ff

[21] H. Moseley, Phil. Mag. **26**, 1024 (1913), 703 (1914)

[22] A. Dauvillier, Compt. Rend. Acad. Sci. **174**,1347 (1922); G. Urbain, ibid p. 1349 [23] http://www-ssrl.slac.stanford.edu/sbsummerschool/studentsonly/ Brown 2005 environ science.pdf;

[24] D. T. Attwood http://www-ssrl.slac.stanford.edu/sbsummerschool/studentsonly/

[25] D. T. Attwood, *Soft X-rays and extreme ultaviolett radiation principles and applications*, Cambridge Univ. Press 2000

[26] W. Mehlhorn, J. Electr. Spectr. Relat. Phenom. **93**,1-15(1998) [27] M. F. Chung and J. H. Jenkins, Surf. Sci., **22**, 479-485 (1970)

[28] P. A.Bennett, J. C.Fuggle, F.U. Hillebrecht, A. Lenselink, and G. A. Sawatzky, Phys. Rev. **B 27**,2194(1983)

[29] B. Lengeler, in: [7]

[30] Alojz Kodre, Iztok Arčon, and Jana Padežnik Gomilšek, Acta Chim. Slov., **51**, 1-10, (2004)

[31] De Fanis A, Saito N, Yoshida H, Senba Y, Tamenori Y, Ohashi H, Tanaka H and Ueda K, Phys. Rev. Lett. **89**, 243001 (2002); K. Ueda J.Phys.B: At. Mol. Opt. Phys. **36**, R1 (2003)

[32] U. Fano, Phys. Rev. **124**, 1866 (1961); U. Fano and J. W. Cooper, Phys. Rev. **A 137**, A1364 (1965)

[33] M. Oura, Y. Tamenori, T. Hayaishi, Y. Kanai, H. Yoshii, K. Tsukamoto, and F. Koike, Phys. Rev **A 70** 022710 (2004)

[34] M. Kato, Y. Morishita, F. Koike, S. Fritzsche, H. Yamaoka, Y. Tamenori, K. Okada, T. Matsudo, T. Gejo, I. H. Suzuki, and N. Saito, J. Phys. B: At. Mol. Opt. Phys. **39** 2059-2069 (2006)

[35] C.T. Chen, Y. Ma, F.Sette,Phys. Rev. **A 40**, 6737 (1989)

[36] figure taken from ADVANCED LIGHT SOURCE Activity Report 2001, p44 see also: W.C. Stolte et al., Phys. Rev. Lett. **86**, 4504 (2001)

[37] J.D. Dehmer and D. Dill, J. Chem. Phys. **65**, 5327 (1976)

[38] S.R. Bare, from: w3.rz-berlin.mpg.de/~ jentoft/lehre/bare_xas_271006.pdf

[39] Matthew Newville, *Fundamentals of XAFS*, University of Chicago, http://cars9.uchicago.edu/~newville/

[40] K. M. Kemner, S. D. Kelly, K. A. Orlandini, A. I. Tsapin, M. G. Goldfeld, Y. D. Perfiliev, and K. H. Nealson, J. Synchrotron Rad. **8**, 949 (2001)

[41] J.G. Chen, Surface Science Reports **30**, 1-152 (1997)

[42] J. Stöhr, K. Baberschke, R.Jaeger, R. Treichler, S. Brennan, Phys. Rev. Lett. **47**, 381 (1981)

[43] E. A. Stern, in: [4]

[44] J.J. Rehr, R.C. Albers, Rev. Mod. Phys. **72**, 621 (2000)

[45] J.E. Penner-Hahn, Coord. Chem. Rev. **190-192**, 1101 (1999)

[46] G. Schütz, W. Wagner, W. Wilhelm, P. Kienle, R. Zeller, R. Frahm, G. Materlik, Phys. Rev. Lett. **58**, 737 (1987)

[47] J. Stöhr, R. Nakajima, IBM J. Res. Develop. **42**, 73(1998) [48] J. Stöhr, J. Electr. Spectr.

Relat. Phenom. **75**, 253 (1995)

[49] J. Stöhr, H. C. Siegmann, *Magnetism – From Fundamentals to Nanoscale Dynamics*, Springer Verlag, Berlin, Heidelberg, New York (2006)

[50] J. Goulon, C. Goulon-Ginet, A. Rogalev, V. Gotte, C. Malgrange, C. Brouder, C. R. Natoli, J. Chem. Phys. **108**, 6394 (1998)

[51] T. Funk, A. Deb, S. J. George, H. Wang, S. P. Cramer, Coord. Chem. Rev. **249**, 3 (2005) [52] C. T. Chen, Y. U. Idzerda, H.-J. Lin, N. V. Smith, G. Meigs, G. H. Ho, E. Pellegrin, and F. Sette, Phys. Rev. Lett. **75**, 152 (1995)

[53] W. Kuch, *Abbildende magnetische Mikrospektroskopie*, Habilitationsschrift, Martin-Luther-Universität Halle-Wittenberg(2002), http://sundoc.bibliothek.uni-halle.de/habil- online/02/03H047/index.htm

[54] U. Fano, Phys. Rev. **178**, 131(1969); addendum Phys. Rev. **184**, 250 (1969) [55] D. T. Pierce, F. Meier, Phys. Rev. Lett. **13**, 5484 (1976)

[56] J. Kessler, *Polarized Electrons*, Springer Verlag, Berlin (1985)

[57] D. T. Pierce, *Atomic, Molecular and Optical Physics: Charged Particles*, Academic Press

(1995) 1

[58] G. F. Hanne, *Polarisierte Elektronen*, www.uni-muenster.de/Physik.PI/Hanne/skripte.html] [59] F. Baudelet, in: [7]

[60] Wolfgang Kuch, Physica Scripta T**109**, 89-95, 2004

[61] W. Kuch, *Imaging Magnetic Microspectroscopy*, in: H. Hopster and H. P. Oepen (Eds.), *Magnetic Microscopy of Nanostructures*, p. 1-28, Springer, Berlin (2004). ISBN 3-540-40186-5, http://www.mpi-halle.mpg.de/mpi/publi/ki/2004/abs pdf/5587 04.html

[62] C. M. Schneider, D. Venus, and J. Kirschner, Phys. Rev. B 45, 5041 (1992)

[63] B. T. Thole, P. Carra, F. Sette und G. van der Laan, Phys. Rev. Lett. **68**, 1943 (1992). [64] P. Carra, B. T. Thole, M. Altarelli und X. Wang, Phys. Rev. Lett. **70**, 694 (1993).

[65] J. Stöhr and H. König, Phys. Rev. Lett. **75**, 3748 (1995) [66] M. Altarelli, Phys. Rev. **B 47**, 597 (1993)

[67] T. Jo, J. Phys. Soc. Japan **62**, 1814 (1993)

[68] A. Ankudinov, J. J. Rehr, Phys. Rev. **B 51**, 1282 (1995)

[69] M. Sacchi and J. Vogel, p. 87 in: E. Beaurepaire, F Scheurer, G. Krill, and J.-P. Kappler (Eds.), *Magnetism and Synchrotron Radiation*, Springer-Verlag, Berlin-Heidelberg (2001)

[70] H. Ebert, Rep. Prog. Phys. **59**, 1665 (1996)

[71] P. Bruno, Phys. Rev. **B 39**, 865 (1989)

[72] A. Dallmeyer, C. Carbone, W. Eberhardt, C. Pampuch, O. Rader, W. Gudat,P. Gambardella, and K. Kern, Phys. Rev. **B 61**, R5133 (2000)

[73] P. Gambardella, S. S. Dhesi, S. Gardonio, C. Grazioli, P. Ohresser, and C. Carbone, Phys. Rev.Lett. **88**, 047202 (2002)

[74] P. Gambardella, S. Rusponi, M. Veronese, S. S. Dhesi, C. Grazioli, A. Dallmeyer, I. Cabria, R. Zeller, P. H. Dederichs, K. Kern, C. Carbone, H. Brune, Science **300**, 1130 (2003)

[75] I. Cabria, B. Nonas, R. Zeller, and P.H. Dederichs, Phys. Rev. **B 65** 054414 (2002)

[76] P. Gambardella, S. Rusponi, M. Veronese, S. S. Dhesi, C. Grazioli, A. Dallmeyer, I. Cabria, R. Zeller, P. H. Dederichs, K. Kern, C. Carbone, H. Brune, Science **300**, 1130 (2003)

[77] J.T. Lau, A. Föhlisch, R. Nietubyć, M. Reif, and W. Wurth, Phys. Rev. Lett. **89** 057201

(2002)

[78] J .T. Lau, A. Föhlisch, M. Martins, R. Nietubyć, M. Reif and W. Wurth, New Journal of Physics **4**, 98.1 (2002)

[79] P.S. Bechthold, in: *Magnetism goes Nano*, 36. IFF-Ferienkurs, Forschungszentrum Jülich (2005), ISBN 3-89336-381-5;
www.fz-juelich.de/iff/src/fs/2005/download/C6 Bechthold FS05.pdf

[80] B. T. Thole, G. van der Laan, and G. A. Sawatzky, Phys. Rev. Lett. **55**, 2086 (1985)

[81] G. van der Laan, B.T. Thole, G.A. Sawatzky, J.B. Goedkoop, J.C. Fuggle, J.-M. Esteva, R. Karnatak, J.P. Remeika, H.A. Dabkowska, Phys. Rev **B 34**, 6529 (1986)

[82] P. Curie, Journal de Physique **3**, 401 (1894) [83] F. de Groot, and J. Vogel, in: [7]

[84] J. Stöhr, www-ssrl.slac.stanford.edu/stohr/xmcd.htm

[85] D. Alders, J. Vogel, C. Levelut, S. D. Peacor, T. Hibma, M. Sacci, L. H. Tjeng, C. T. Chen, G. van der Laan, B. T. Thole, and G. A. Sawatzky, Europhys. Lett. **32**, 259 (1995).

[86] figure taken from: J. Stöhr, S. Anders, IBM J. Res. Develop. **44**, 535 (2000) [87] J. L.

Erskine and E. A. Stern, Phys. Rev. **B 12**, 5016 (1975)

[88] B.W. Shore and D.H. Menzel, *Principles of Atomic Spectra*, John Wiley and Sons, New York (1968)

[89] A. Messiah, *Quantum Mechanics, Vol. II*, North Holland Publishing Company, Amsterdam (1970)

[90] H.A. Bethe, E.E. Salpeter, *Quantum Mechanics of One and Two Electron Atoms*, Plenum, New York (1977)

Index

The IFF Spring School & Scattering Methods in Jülich

PGI-JCNS-TA
52425 Jülich
Phone: *++49 2461 61-1739*
Fax: *++49 2461 61-2410*
email: *springschool@fz-juelich.de*
web: *www.iff-springschool.de*

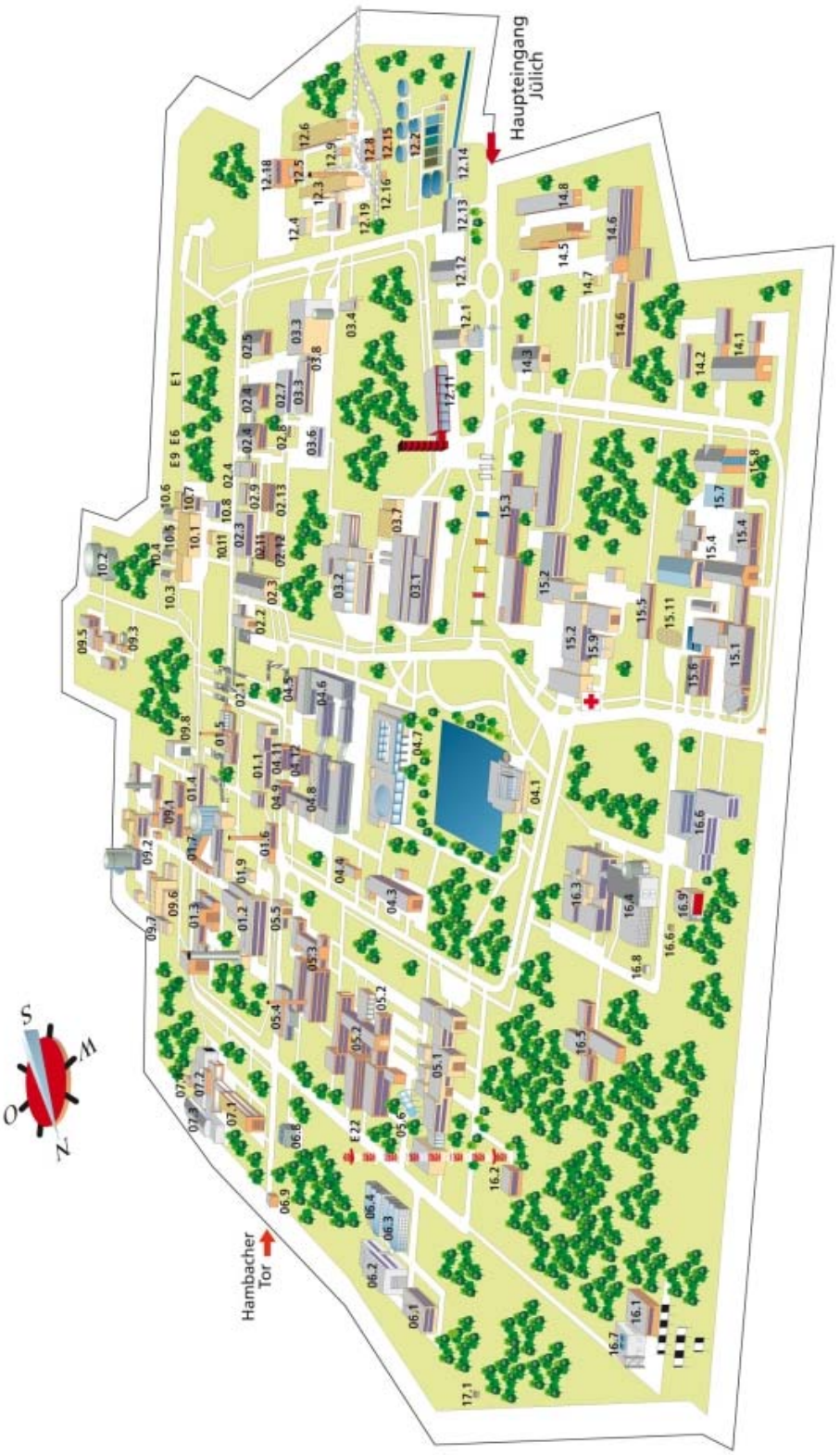
The IFF Spring Schools were first brought into being in 1970 by the Jülich "Institute of Solid State Research" (IFF). Since then, the schools have made it possible for students and young scientists to gain a two-week insight into a current topic related to condensed matter physics. In 2011, IFF was dissolved as part of a restructuring process within Forschungszentrum Jülich, and new institutes, namely the "Peter Grünberg Institute" (PGI), the "Jülich Centre for Neutron Science" (JCNS) and the "Institute of Complex Systems" (ICS) were established. Together with the "Institute for Advanced Simulation" (IAS), the three institutes continue to organize the IFF Spring Schools. The 43rd Spring School 2012 was organized jointly by ICS-1/JCNS-1 "Neutron Scattering" and PGI-4/JCNS-2 "Scattering Methods".

The Institute for Neutron Scattering conducts neutron research, focusing mainly on soft matter. This term is used to indicate materials which respond strongly to weak forces. Typical examples are rubber materials and emulsions. In addition to their structure, the dynamics of these systems are also of interest. Neutron scattering allows us to investigate both of these aspects simultaneously. The institute has modern chemical laboratories at its disposal in order to produce or alter soft matter. As well as neutron scattering instruments, we make use of various standard methods in order to study and characterize soft matter.

At the Institute for Scattering Methods, work focuses on the investigation of structural and magnetic orders, fluctuations and excitations in complex or nanostructured magnetic systems and highly correlated electron systems with the aid of the most advanced synchrotron X-ray and neutron scattering methods. Our aim is to relate this microscopic information to macroscopic physical properties and functionalities. The systems being investigated have huge potential for future applications in information technology. A further aspect of our work lies in preparing samples and sample characterization.

The development of novel scattering techniques and the construction and continuous improvement of instruments on large scale devices are the main focal points of the work of both institutes that together make up the "Jülich Centre for Neutron Science". JCNS operates neutron scattering instruments at leading facilities in Europe and the USA, and makes them accessible to external users via a peer-review system.

Forschungszentrum Jülich - Campus Map



1. **Soft Matter**
   From Synthetic to Biological Materials
   Lecture manuscripts of the 39th IFF Spring School March 3 – 14, 2008
   Jülich, Germany
   edited by J.K.G. Dhont, G. Gompper, G. Nägele, D. Richter, R.G. Winkler (2008),
   c. 1000 pages
   ISBN: 978-3-89336-517-3

2. **Structural analysis of diblock copolymer nanotemplates using grazing incidence scattering**
   by D. Korolkov (2008), III, 167 pages
   ISBN: 978-3-89336-522-7

3. **Thermal Nonequilibrium**
   Thermal forces in fluid mixtures
   Lecture Notes of the 8th International Meeting on Thermodiffusion,
   9 – 13 June 2008, Bonn, Germany
   edited by S. Wiegand, W. Köhler (2008), 300 pages
   ISBN: 978-3-89336-523-4

4. **Synthesis of CMR manganites and ordering phenomena in complex transition metal oxides**
   by H. Li (2008), IV, 176 pages
   ISBN: 978-3-89336-527-2

5. **Neutron Scattering**
   Lectures of the JCNS Laboratory Course held at the Forschungszentrum Jülich and the research reactor FRM II of TU Munich
   edited by R. Zorn, Th. Brückel, D. Richter (2008), ca. 500 pages
   ISBN: 978-3-89336-532-6

6. **Ultrafast Magnetization Dynamics**
   by S. Woodford (2008), 130 pages
   ISBN: 978-3-89336-536-4

7. **Role of Surface Roughness inTribology: from Atomic to Macroscopic Scale**
   by C. Yang (2008), VII, 166 pages
   ISBN: 978-3-89336-537-1

8. **Strahl- und Spindynamik von Hadronenstrahlen in Mittelenergie-Ringbeschleunigern**
   von A. Lehrach (2008), II, 171 Seiten
   ISBN: 978-3-89336-548-7

9. **Phase Behaviour of Proteins and Colloid-Polymer Mixtures**
   by C. Gögelein (2008), II, 147 pages
   ISBN: 978-3-89336-555-5

10. **Spintronics – From GMR to Quantum Information**
    Lecture Notes of the 40th IFF Spring School March 9 – 20, 2009
    Jülich, Germany
    edited by St. Blügel, D. Bürgler, M. Morgenstern, C. M. Schneider,
    R. Waser (2009), c. 1000 pages
    ISBN: 978-3-89336-559-3

11. **ANKE / PAX Workshop on SPIN Physics**
    JINR, Dubna, Russia / June 22. – 26, 2009
    Org. Committee: A. Kacharava, V. Komarov, A. Kulikov, P. Lenisa, R. Rathmann,
    H. Ströher (2009), CD-ROM
    ISBN: 978-3-89336-586-9

12. **Entwicklung einer Nanotechnologie-Plattform für die Herstellung Crossbar-basierter Speicherarchitekturen**
    von M. Meier (2009), 135 Seiten
    ISBN: 978-3-89336-598-2

13. **Electronic Oxides –**
    Correlation Phenomena, Exotic Phases and Novel Functionalities
    Lecture Notes of the 41st IFF Spring School March 8 – 19, 2010
    Jülich, Germany
    edited by St. Blügel, T. Brückel, R. Waser, C.M. Schneider (2010), ca. 1000 pages
    ISBN: 978-3-89336-609-5

14. **4th Georgian-German School and Workshop in Basic Science**
    Tbilisi, Georgia / May 3 – 7, 2010
    Org. Committee: E. Abrosimova, R. Botchorishvili, A. Kacharava, M. Nioradze,
    A. Prangishvili, H. Ströher (2010); CD-ROM
    ISBN: 978-3-89336-629-3

15. **Neutron Scattering**
    Lectures of the JCNS Laboratory Course held at Forschungszentrum Jülich and the research reactor FRM II of TU Munich
    edited by Th. Brückel, G. Heger, D. Richter, G. Roth and R. Zorn (2010),
    ca 350 pages
    ISBN: 978-3-89336-635-4

16. **Ab *initio* investigations of magnetic properties of ultrathin transition-metal films on *4d* substrates**
    by A. Al-Zubi (2010), II, 143 pages
    ISBN: 978-3-89336-641-5

17. **Investigation of a metal-organic interface realization and understanding of a molecular switch**
    by O. Neucheva (2010), 134 pages
    ISBN: 978-3-89336-650-7

18. **Reine Spinströme in lateralen Spinventilen, *in situ* Erzeugung und Nachweis**
von J. Mennig (2010), V, 95 Seiten
ISBN: 978-3-89336-684-2

19. **Nanoimprint Lithographie als Methode zur chemischen Oberflächenstrukturierung für Anwendungen in der Bioelektronik**
von S. Gilles (2010), II, 169 Seiten
ISBN: 978-3-89336-686-6

20. **Macromolecular Systems in Soft- and Living-Matter**
Lecture Notes of the 42nd IFF Spring School 2011 February 14 – 25, 2011 Jülich, Germany
edited by J. K.G. Dhont, G. Gompper, P. R.Lang, D. Richter, M. Ripoll, D. Willbold, R. Zorn (2011), ca. 1000 pages
ISBN: 978-3-89336-688-0

21. **The spin structure of magnetic nanoparticles and in magnetic nanostructures**
by S. Disch (2011), V, 342 pages
ISBN: 978-3-89336-704-7

22. **Element-selective and time-resolved magnetic investigations in the extreme ultraviolet range**
by P. Grychtol (2011), xii, 144 pages
ISBN: 978-3-89336-706-1

23. **Spin-Transfer Torque Induced Dynamics of Magnetic Vortices in Nanopillars**
by V. Sluka (2011), 121 pages
ISBN: 978-3-89336-717-7

24. **Adsorption von Phthalocyaninen auf Edelmetalloberflächen**
von I. Kröger (2011), vi, 206 Seiten
ISBN: 978-3-89336-720-7

25. **Time-Resolved Single Molecule FRET Studies on Folding/Unfolding Transitions and on Functional Conformational Changes of Phosphoglycerate Kinase**
by T. Rosenkranz (2011), III, 139 pages
ISBN: 978-3-89336-721-4

26. **NMR solution structures of the MloK1 cyclic nucleotide-gated ion channel binding domain**
by S. Schünke (2011), VI, (getr. pag.)
ISBN: 978-3-89336-722-1

27. **Neutron Scattering**
    Lectures of the JCNS Laboratory Course held at Forschungszentrum Jülich and the research reactor FRM II of TU Munich
    edited by Th. Brückel, G. Heger, D. Richter, G. Roth and R. Zorn (2011), ca 350 pages
    ISBN: 978-3-89336-725-2

28. **Neutron Scattering**
    Experiment Manuals of the JCNS Laborator Course held at Forschungszentrum Jülich and the research reactorFRM II of TU Munich
    edited by Th. Brückel, G. Heger, D. Richter, G. Roth and R. Zorn (2011), ca. 180 pages
    ISBN: 978-3-89336-726-9

29. **Silicon nanowire transistor arrays for biomolecular detection**
    by X.T.Vu (2011), vii, 174 pages
    ISBN: 978-3-89336-739-9

30. **Interactions between parallel carbon nanotube quantum dots**
    by K. Goß (2011), viii, 139 pages
    ISBN: 978-3-89336-740-5

31. **Effect of spin-orbit scattering on transport properties of low-dimensional dilute alloys**
    by S. Heers (2011), viii, 216 pages
    ISBN: 978-3-89336-747-4

32. **Charged colloids and proteins: Structure, diffusion, and rheology**
    by M. Heinen (2011), xii, 186 pages
    ISBN: 978-3-89336-751-1

33. **Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources**
    Lecture Notes of the 43rd IFF Spring School 2012
    March 5 – 16, 2012 Jülich, Germany
    edited by M. Angst, T. Brückel, D. Richter, R. Zorn ca. 1000 pages
    ISBN: 978-3-89336-759-7